

⑫ 公開特許公報(A)

昭63-191372

⑮ Int.Cl.⁴

G 11 B 20/12

識別記号

庁内整理番号

8524-5D

⑬ 公開 昭和63年(1988)8月8日

審査請求 有 発明の数 6 (全21頁)

⑭ 発明の名称 情報記録媒体および情報再生方式

⑯ 特 願 昭62-219962

⑰ 出 願 昭62(1987)9月2日

優先権主張 ⑱ 昭61(1986)9月3日 ⑲ 日本(JP) ⑳ 特願 昭61-207591

㉑ 発 明 者 関 口 博 司 東京都千代田区外神田1丁目9番9号 内田ビル5F カ
ナース・データー株式会社㉒ 出 願 人 カナース・データー株 東京都千代田区外神田1丁目9番9号 内田ビル5F
式会社

㉓ 代 理 人 弁理士 長谷川 芳樹 外1名

明 細 書

1. 発明の名称

情報記録媒体および情報再生方式

2. 特許請求の範囲

1. 音声情報が記録された情報記録媒体において、

複数の区画に区分した第1音声情報列が記録された第1の領域と、

前記第1音声情報列に対応した内容でかつ別の音声情報からなる第2音声情報列が、前記第1音声情報列の区画に対応する複数の区画に区分して記録された第2の領域と、

前記第1および第2音声情報列のそれぞれの記録位置を前記区画ごとに示す記録位置識別情報が記録された第3の領域と

を少なくとも備えることを特徴とする情報記録媒体。

2. 前記第3の領域は、前記記録位置識別情

報の少なくともその一部がディレクトリとして所定の位置にひとまとめに記録された領域である特許請求の範囲第1項記載の情報記録媒体。

3. 前記第3の領域は、前記第1および第2音声情報列のそれぞれの記録位置に近接した位置に識別情報として前記記録位置識別情報の少なくともその一部が記録された領域である特許請求の範囲第1項記載の情報記録媒体。

4. 前記第1および第2の領域における前記区画は1又は2以上の小区画からなり、

前記第3の領域は前記小区画ごとの記録位置識別情報をも包含している特許請求の範囲第1項記載の情報記録媒体。

5. 前記情報記録媒体はCD等の円板型記録媒体である特許請求の範囲第1項記載の情報記録媒体。

6. 前記情報記録媒体はDAT等のテープ型記録媒体である特許請求の範囲第1項記載の情報記録媒体。

7. 音声情報が記録された情報記録媒体にお

いて、

複数の区画に区分した第1音声情報列が記録された第1の領域と、

前記第1音声情報列に対応した内容でかつ別の音声情報からなる第2音声情報列が前記第1音声情報列の区画に対応する複数の区画に区分して記録された第2の領域と、

前記第1および第2音声情報列に対応した内容でかつ別の音声情報からなる第3音声情報列が前記区画の1又は2以上をひとまとまりとした複数の区画群に区分して記録された第3の領域と、

前記第1および第2音声情報列のそれぞれの記録位置を前記区画ごとに示す記録位置識別情報が記録され、かつ前記第3音声情報列の記録位置を前記区画群ごとに示す記録位置識別情報が記録された第4の領域と

を少なくとも備えることを特徴とする情報記録媒体。

8. 前記第4の領域は、前記記録位置識別情報の少なくともその一部がディレクトリとして所

複数の区画に区分した第1音声情報列が記録された第1の領域と、

前記第1音声情報列に対応した内容でかつ別の音声情報からなる第2音声情報列が前記第1音声情報列の区画に対応する複数の区画に区分して記録された第2の領域と、

前記第1又は第2音声情報列に対応する内容の文字情報からなる文字情報列が、前記第1音声情報列の区画に対応する複数の区画に区分して記録された第3の領域と、

前記第1、第2音声情報列および文字情報列のそれぞれの記録位置を前記区画ごとに示す記録位置識別情報が記録された第4の領域と

を少なくとも備えることを特徴とする情報記録媒体。

14. 前記第4の領域は、前記記録位置識別情報の少なくともその一部がディレクトリとして所定の位置にひとまとめに記録された領域である特許請求の範囲第13項記載の情報記録媒体。

15. 前記第4の領域は、前記第1および第

定の位置にひとまとめに記録された領域である特許請求の範囲第7項記載の情報記録媒体。

9. 前記第4の領域は、前記第1、第2および第3音声情報列のそれぞれの記録位置に近接した位置に識別情報として前記記録位置識別情報の少なくともその一部が記録された領域である特許請求の範囲第7項記載の情報記録媒体。

10. 前記第1および第2の領域における前記区画は1又は2以上の小区画からなり、

前記第4の領域は前記小区画ごとの記録位置識別情報を記録している特許請求の範囲第7項記載の情報記録媒体。

11. 前記情報記録媒体はCD等の円板型記録媒体である特許請求の範囲第7項記載の情報記録媒体。

12. 前記情報記録媒体はDAT等のテープ型記録媒体である特許請求の範囲第7項記載の情報記録媒体。

13. 音声情報および文字情報が記録された情報記録媒体において、

2音声情報列および文字情報列のそれぞれの記録位置に近接した位置に識別情報として前記記録位置識別情報の少なくともその一部が記録された領域である特許請求の範囲第13項記載の情報記録媒体。

16. 前記第1、第2および第3の領域における前記区画は1又は2以上の小区画からなり、

前記第4の領域は前記小区画ごとの記録位置識別情報が記録された領域である特許請求の範囲第13項記載の情報記録媒体。

17. 前記情報記録媒体はCD等の円板型記録媒体である特許請求の範囲第13項記載の情報記録媒体。

18. 前記情報記録媒体はDAT等のテープ型記録媒体である特許請求の範囲第13項記載の情報記録媒体。

19. 媒体にあらかじめ記録されている音声情報を再生する情報再生方式において、

前記音声情報は、複数の区画に区分した第1音声情報列と、前記第1音声情報列に対応した内容

でかつ別の音声情報からなり前記第1音声情報列の区画に対応する複数の区画に区分した第2音声情報列と、前記第1および第2音声情報列のそれぞれの前記媒体における記録位置を前記区画ごとに示す記録位置識別情報とを少なくとも含み、

前記第1音声情報列の再生中または中断の後に前記第2音声情報列の再生指令が入力されたときは再生中の前記第1音声情報列中の区画に対応する前記第2音声情報列中の区画の音声情報を前記記録位置識別情報にもとづいて前記媒体から読み出し再生する第1のステップと、

前記第2音声情報列の再生中または中断の後に前記第1音声情報列の再生指令が入力されたときは再生中の前記第2音声情報列中の区画に対応する前記第1音声情報列中の区画の音声情報を前記記録位置識別情報にもとづいて前記媒体から読み出し再生する第2のステップと

を少なくとも備えることを特徴とする情報再生方式。

20. 前記第1のステップは、前記再生指令

に区分した第3音声情報列と、前記第1および第2音声情報列のそれぞれの前記媒体における記録位置を前記区画ごとに示し、かつ前記第3音声情報列の前記媒体における記録位置を前記区画群ごとに示す記録位置識別情報とを少なくとも含み、

前記第1音声情報列の再生中または中断の後に前記第2音声情報列の再生指令が入力されたときは再生中の前記第1音声情報列中の区画に対応する前記第2音声情報列中の区画の音声情報を前記記録位置識別情報にもとづいて前記媒体から読み出し再生する第1のステップと、

前記第2音声情報列の再生中または中断の後に前記第1音声情報列の再生指令が入力されたときは再生中の前記第2音声情報列中の区画に対応する前記第1音声情報列中の区画の音声情報を前記記録位置識別情報にもとづいて前記媒体から読み出し再生する第2のステップと、

前記第1又は第2音声情報列の再生中または中断の後に前記第3音声情報列の再生指令が入力されたときは再生中の前記第1又は第2音声情報列

の入力時に戻し指令が入力されたときは指令された量だけ前記媒体からの前記第2音声情報列中の音声情報の読み出し位置を前記区画に関連したきざみで戻すステップであることを特徴とする特許請求の範囲第19項記載の情報再生方式。

21. 前記第2のステップは、前記再生指令の入力時に戻し指令が入力されたときは指令された量だけ前記媒体からの前記第1音声情報列中の音声情報の読み出し位置を前記区画に関連したきざみで戻すステップであることを特徴とする特許請求の範囲第19項記載の情報再生方式。

22. 媒体にあらかじめ記録されている音声情報を再生する情報再生方式において、

前記音声情報は、複数の区画に区分した第1音声情報列と、前記第1音声情報列に対応した内容でかつ別の音声情報からなり前記第1音声情報列の区画に対応する複数の区画に区分した第2音声情報列と、前記第1および第2音声情報列に対応した内容でかつ別の音声情報からなり前記区画の1又は2以上をひとまとまりとした複数の区画群

中の区画に対応する前記第3音声情報列中の区画群の音声情報を前記記録位置識別情報にもとづいて前記媒体から読み出し再生する第3のステップと

を少なくとも備えることを特徴とする情報再生方式。

23. 前記第1のステップは、前記再生指令の入力時に戻し指令が入力されたときは指令された量だけ前記媒体からの前記第2音声情報列中の音声情報の読み出し位置を前記区画に関連したきざみで戻すステップであることを特徴とする特許請求の範囲第22項記載の情報再生方式。

24. 前記第2のステップは、前記再生指令の入力時に戻し指令が入力されたときは指令された量だけ前記媒体からの前記第1音声情報列中の音声情報の読み出し位置を前記区画に関連したきざみで戻すステップであることを特徴とする特許請求の範囲第22項記載の情報再生方式。

25. 前記第3のステップは、前記再生指令の入力時に戻し指令が入力されたときは指令され

た量だけ前記媒体からの前記第3音声情報列中の音声情報の読み出し位置を前記区画群に関連したきざみで戻すステップであることを特徴とする特許請求の範囲第22項記載の情報再生方式。

26. 媒体にあらかじめ記録されている音声情報および文字情報を再生する情報再生方式において、

前記音声情報および文字情報は、複数の区画に区分した第1音声情報列と、前記第1音声情報列に対応した内容でかつ別の音声情報からなり前記第1音声情報列の区画に対応する複数の区画に区分した第2音声情報列と、前記第1又は第2音声情報列に対応する内容の文字情報からなり前記第1音声情報列の区画に対応する複数の区画に区分した文字情報列と、前記第1、第2音声情報列および文字情報列のそれぞれの前記媒体における記録位置を前記区画ごとに示す記録位置識別情報とを少なくとも含み、

前記第1音声情報列の再生中または中断の後に前記第2音声情報列の再生指令が入力されたとき

は再生中の前記第1音声情報列中の区画に対応する前記第2音声情報列中の区画の音声情報を前記記録位置識別情報にもとづいて前記媒体から読み出し再生する第1のステップと、

前記第2音声情報列の再生中または中断の後に前記第1音声情報列の再生指令が入力されたときは再生中の前記第2音声情報列中の区画に対応する前記第1音声情報列中の区画の音声情報を前記記録位置識別情報にもとづいて前記媒体から読み出し再生する第2のステップと、

前記第1又は第2音声情報列の再生中または中断の後は前記文字情報列中の対応する区画の文字情報を前記記録位置識別情報にもとづいて前記媒体から読み出しディスプレイ表示する第3のステップと

を少なくとも備えることを特徴とする情報再生方式。

27. 前記第1のステップは、前記再生指令の入力時に戻し指令が入力されたときは指令された量だけ前記媒体からの前記第2音声情報列中の

音声情報の読み出し位置を前記区画に関連したきざみで戻すステップであることを特徴とする特許請求の範囲第26項記載の情報再生方式。

28. 前記第2のステップは、前記再生指令の入力時に戻し指令が入力されたときは指令された量だけ前記媒体からの前記第1音声情報列中の音声情報の読み出し位置を前記区画に関連したきざみで戻すステップであることを特徴とする特許請求の範囲第26項記載の情報再生方式。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明はCD-ROM、DAT等の媒体に音声情報等を記録した情報記録媒体と、このような媒体にあらかじめ記録された音声情報等を再生する情報再生方式に関する。

〔従来の技術〕

英会話等の語学の独習用、詩吟の練習用、法律の独習用、その他の目的のために、カセットテープ等の媒体に音声情報を記録したものが種々提供

されている。ここで英会話の独習用のものを例に説明すると、従来のものは例えば、一連の英語の発声（音声情報）が記録されているカセットテープ（又はレコード）であり、このテープ教材とテキストの組み合わせで独習者に使われるようになっている。そして、そのレベルは初級用から上級用まで種々のものがある。

従来のものでも、テープを聴いているときに聴き取れないところが出てきたら、テープを少し巻き戻して再度聴き直すことはできる。また、何度も繰り返して聴くことも可能である。しかし、たとえ何度聴き直しても、聴き取れない部分があくまで聴き取れないことは、外国語の勉強にはつきものである。そのような時、従来のものではテキストを参照する以外に方法がなかった。なぜなら、例えば米語のように複数の単語を一個の単語のように発音している場合には、テープをゆっくり回して聴いても、聴き取れないことが多いからである。

ところが、上述の如くテープでは聴き取れない

部分があるときに、これをテキストを参照してから理解することについては、以下の如く主として2つの問題がある。

第1の問題は、現代の日本人は読解力に比べヒヤリング能力が極端に低いので、文字にたよる癖がついており、そのため、いつまでたってもヒヤリング能力が向上しない傾向があることである。ヒヤリングの訓練中はなるべく文字を見ないことが望ましい。ところが、従来方式のテープ（又はその他の媒体）を使った練習システムでは、上記のような場面に出会ったときはテキストを見る以外に適切な方法がない。テキストを見る必要を感じるのが5分あるいは10分に1回程度であれば、このような弊害も少ないが、逆にこのような教材はその人に易しすぎることであり訓練にならない。訓練に適したものを選ぶと、頻りに聴きとれないところが出てくる。その時、テキストを見ていたのでは文字にたよる悪癖から抜け出せないことになる。

第2の問題は、必要を感じるごとにテキストを

見るのがわずらわしいことである。特に、電車の中などでヒヤリング訓練をしている時は、テキストを見なくても済ませるようにすることが望ましい。

そこで、従来の技術によって上記のテキストを参照しなければならないという問題を解決しようとした場合には、次の2つの方式が考えられる。

第1は、ネイティブスピーカーが話す自然な速さの英語の音声情報と、これを日本人が聴き取り易くするためにゆっくりと単語を区切って話した英語の音声情報とをシーケンシャルに記録し、更に必要に応じてこれら英語の音声情報を日本語で解説した日本語の音声情報をシーケンシャルに記録する方式である。具体的に説明すると、まずネイティブスピーカーの話す速さで

"It's not much of a problem. I'd second that."

を1.8秒の長さで記録し、次にゆっくりとした速さで

"It is not much of a problem. I would

second that."

を5.0秒の長さで記録し、次に日本語で

「"It"とは前の〇〇を指しており、"not much of a problem"は慣用句で『問題ない』とか『心配ない』というような意味です。次に、

"I'd"とは"I would"をつめた言い方であり、同じように"I could"も"I'd"とつめて言えます。"second"とは『指示する』、『賛成する』という意味があり、"I'd second that"で『同感です』という意味になります。』

を3.7秒の長さで記録する。

この記録はシーケンシャルになされており、従って操作者がそれぞれの頭出しをしない限りシーケンシャルに再生されていく。

第2の方式は、上記のネイティブスピーカーの話す自然な速さの英語の音声情報を例えばテープの第1のトラックに記録し、ゆっくりと単語を区切って話す英語の音声情報を第2のトラックに記録し、更に必要に応じて日本語の解説を他のトラックに記録しておく方式である。この方式によれば、

再生するトラックを指定すれば、希望する内容の音声情報を再生することができる。

〔発明が解決しようとする問題点〕

ところが、英会話テープを使う人には上級者もいれば初級者もいる。従って、上記第1の方式によれば、下記の如き問題がある。すなわち、上級者にとってはネイティブスピーカーの話す英語を聴くことが主目的となり、従ってその間に入ってくるゆっくりと話す英語や日本語解説を聴くことはわずらわしくなる。そこで、ネイティブスピーカーの話す英語の部分のみを選択して再生しようとしても、記録がシーケンシャルにされているため頭出しは容易ではなく、複雑なスイッチ操作が必要になる。

他方、初級者にとってはネイティブスピーカーの話す英語は速すぎて聴き取ることができず、そこでゆっくり話す英語と日本語解説を対比して聴きたくなる。ところが、第1の方式のようにこれらが順次聴こえてくるのでは細部について十分に理解できず、また文章が長いときや聴き取れなかつ

た部分が複数あるときにはシーケンシャルに聴いただけでは理解できない。そこで、ネイティブスピーカの話す部分と、ゆっくり単語を区切って話す部分と、必要に応じて日本語解説の部分とを対比して聴こうとしても、記録はシーケンシャルにされているため頭出しは容易でなく、極めて複雑なスイッチ操作が必要になる。

これに対して上記第2の方式によれば、前述の如く各音声情報は異なるトラックに別けて記録されているので、例えばネイティブスピーカの話す部分のみを連続的に再生することは比較的容易である。しかしながら、例えば第1のトラックのネイティブスピーカの英語が聴き取れなかった場合に、自分の聴き取れなかったところに相当する第2のトラック上での記録位置を自分で見つけ出すことは困難である。もちろん、第1のトラックのテープカウンタと、第2のトラックのテープカウンタの対応表を作り、第1のトラックのネイティブスピーカの英語がある記録位置で聴き取れなかったら、その対応表に従い第2のトラックの対応

する記録位置までテープを進めあるいは戻し、ゆっくりと話す英語を聴くこともできる。また、これを自動的にすることも不可能ではない。しかし、テープの往復に時間がかかる等の欠点がある。

このため、従来の英会話独習用の情報記録媒体あるいは情報再生方式では、上級者も下級者も共に満足できるようなものは提供できなかった。このような問題点は、語学独習用のものに限らず詩吟の練習用のもの、法律の学習用のもの等にもあった。

そこで本発明は、上級者すなわち慣れた者にも、初級者すなわち慣れていない者にも、適切に音声情報等を与えてこれらの者を共に満足させることのできる情報記録媒体と情報再生方式を提供することを目的とする。

(問題点を解決するための手段)

本発明に係る情報記録媒体の第1の態様は、媒体(例えばCD-ROM、DAT)に音声情報を記録した情報記録媒体であって、次の3つの領域を少なくとも備えることを特徴とする。すなわち、

第1音声情報列に対応した内容でかつ別の音声情報からなる第2音声情報列(例えば単語を区切って話す英語の音声情報列)を第1音声情報列の区画に対応する複数の区画に区分して記録した第2の領域と、

第1および第2音声情報列に対応した内容でかつ別の音声情報からなる第3音声情報列(例えば日本語で解説した音声情報列)を第1および第2音声情報列の区画の1又は2以上をひとまとまりとした複数の区画群に区分して記録した第3の領域と、

第1および第2音声情報列のそれぞれの記録位置を区画ごとに示す記録位置識別情報を例えばディレクトリとして記録し、かつ第3音声情報列の記録位置を区画群ごとに示す記録位置識別情報を例えばディレクトリとして記録した第4の領域との4領域を少なくとも備える。

本発明に係る情報記録媒体の第3の態様は、媒体(例えばCD-ROM、DAT)に音声情報を記録した情報記録媒体であって、次の4つの領域

複数の区画に区分した第1音声情報列(例えばネイティブスピーカの話す英語の音声情報列)を記録した第1の領域と、

第1音声情報列に対応した内容でかつ別の音声情報からなる第2音声情報列(例えば単語を区切って話す英語の音声情報列)を第1音声情報列の複数の区画に対応する複数の区画に区分して記録した第2の領域と、

第1および第2音声情報列のそれぞれの媒体における記録位置を区画ごとに示す記録位置識別情報を例えばディレクトリとして記録した第3の領域と

の3領域を少なくとも備える。

本発明に係る情報記録媒体の第2の態様は、媒体(例えばCD-ROM、DAT)に音声情報を記録した情報記録媒体であって、次の4つの領域を少なくとも備えることを特徴とする。すなわち、

複数の区画に区分した第1音声情報列(例えばネイティブスピーカの話す英語の音声情報列)を媒体に記録する第1のステップと、

を少なくとも備えることを特徴とする。すなわち、
 複数の区画に区分した第1音声情報列（例えばネイティブスピーカの話す英語の音声情報列）を記録した第1の領域と、

第1音声情報列に対応した内容でかつ別の音声情報からなる第2音声情報列（例えば単語を区切って話す英語の音声情報列）を第1音声情報列の区画に対応する複数の区画に区分して記録した第2の領域と、

第1又は第2音声情報列に対応する内容の文字情報からなる文字情報列を第1音声情報列の区画に対応する複数の区画に区分して記録した第3の領域と、

第1、第2音声情報列および文字情報列のそれぞれの記録位置を区画ごとに示す記録位置識別情報を例えばディレクトリとして記録した第4の領域と

の4領域を少なくとも備える。

また、本発明に係る情報再生方式の第1の様態は、媒体（例えばCD-ROM、DAT）にあら

かじめ記録されている音声情報を再生する情報再生方式であって、音声情報は次の3つの情報を少なくとも含み、かつ下記の2つのステップを少なくとも備えることを特徴とする。すなわち、

音声情報は、複数の区画に区分した第1音声情報列（例えばネイティブスピーカの話す英語の音声情報列）と、第1音声情報列に対応した内容でかつ別の音声情報からなり第1音声情報列の複数の区画に対応する複数の区画に区分した第2音声情報列（例えば単語を区切って話す英語の音声情報列）と、第1および第2音声情報列のそれぞれの媒体における記録位置を区画ごとに示す記録位置識別情報（例えばディレクトリとして記録された情報）とを少なくとも含み、

第1音声情報列の再生中または中断の後に第2音声情報列の再生指令が入力されたときは、再生中または中断の後の第1音声情報列中の区画に対応する第2音声情報列中の区画の音声情報を記録位置識別情報にもとづいて媒体から読み出し再生する第1のステップと、

第2音声情報列の再生中または中断の後に第1音声情報列の再生指令が入力されたときは、再生中の第2音声情報列中の区画に対応する第1音声情報列中の区画の音声情報を記録位置識別情報にもとづいて媒体から読み出し再生する第2のステップと

の2ステップを少なくとも備える。

本発明に係る情報再生方式の第2の様態は、媒体（例えばCD-ROM、DAT）にあらかじめ記録されている音声情報を再生する情報再生方式であって、音声情報は次の4つの情報を少なくとも含み、かつ下記の3つのステップを少なくとも備えることを特徴とする。すなわち、

音声情報は、複数の区画に区分した第1音声情報列（例えばネイティブスピーカの話す英語の音声情報列）と、第1音声情報列に対応した内容でかつ別の音声情報からなり第1音声情報列の区画に対応する複数の区画に区分した第2音声情報列（例えば単語を区切って話す英語の音声情報列）と、第1および第2音声情報列に対応した内容で

かつ別の音声情報からなり上記区画の1又は2以上をひとまとまりとした複数の区画群に区分した第3音声情報列（例えば日本語で解説した音声情報列）と、第1および第2音声情報列のそれぞれの媒体における記録位置を区画ごとに示し、かつ第3音声情報列の媒体における記録位置を区画群ごとに示す記録位置識別情報（例えばディレクトリとして記録された情報）とを少なくとも含み、

第1音声情報列の再生中または中断の後に第2音声情報列の再生指令が入力されたときは再生中の第1音声情報列中の区画に対応する第2音声情報列中の区画の音声情報を記録位置識別情報にもとづいて媒体から読み出し再生する第1のステップと、

第2音声情報列の再生中または中断の後に第1音声情報列の再生指令が入力されたときは再生中の第2音声情報列中の区画に対応する第1音声情報列中の区画の音声情報を記録位置識別情報にもとづいて媒体から読み出し再生する第2のステップと、

第1又は第2音声情報列の再生中または中断の後に第3音声情報列の再生指令が入力されたときは再生中の第1又は第2音声情報列中の区画に対応する第3音声情報列中の区画群の音声情報を記録位置識別情報にもとづいて媒体から読み出し再生する第3のステップと

の3ステップを少なくとも備える。

本発明に係る情報再生方式の第3の態様は、媒体(例えばCD-ROM、DAT)にあらかじめ記録されている音声情報および文字情報を再生する情報再生方式であって、音声情報および文字情報は次の4つの情報を少なくとも含み、かつ下記の3つのステップを少なくとも備えることを特徴とする。すなわち、

音声情報および文字情報は、複数の区画に区分した第1音声情報列(例えばネイティブスピーカの話す英語の音声情報列)と、第1音声情報列に対応した内容でかつ別の音声情報からなり第1音声情報列の区画に対応する複数の区画に区分した第2音声情報列(例えば単語を区切って話す英語

中の対応する区画の文字情報を記録位置識別情報にもとづいて媒体から読み出し例えばLCD等でディスプレイ表示する第3のステップと

の3ステップを少なくとも備える。

〔作用〕

本発明に係る情報記録媒体の第1の態様は以上のように構成されるので、第1音声情報列を記録する第1の領域は、基本となる音声情報を例えば発声上の区切り又は言語上なんらかの区切りにもとづき、あるいはこれら区切りと特に係わりなく複数の区画に区分して記録するように働き、第2音声情報列を記録する第2の領域は、第1音声情報列の音声情報を例えば言い換えた音声情報を複数の区画に区分して記録するように働き、記録位置識別情報を記録する第3のステップは、上記2つの情報列の媒体における記録位置を区画ごとに識別させるように働く。

本発明に係る情報記録媒体の第2の態様は以上のように構成されるので、第1音声情報列を記録する第1の領域は、基本となる音声情報を例えば

の音声情報列)と、第1又は第2音声情報列に対応する内容の文字情報からなり第1音声情報列の区画に対応する複数の区画に区分した文字情報列と、第1、第2音声情報列および文字情報列のそれぞれの媒体における記録位置を区画ごとに示す記録位置識別情報(例えばディレクトリとして記録された情報)とを少なくとも含み、

第1音声情報列の再生中または中断の後に第2音声情報列の再生指令が入力されたときは再生中の第1音声情報列中の区画に対応する第2音声情報列中の区画の音声情報を記録位置識別情報にもとづいて媒体から読み出し再生する第1のステップと、

第2音声情報列の再生中または中断の後に第1音声情報列の再生指令が入力されたときは再生中の第2音声情報列中の区画に対応する第1音声情報列中の区画の音声情報を記録位置識別情報にもとづいて媒体から読み出し再生する第2のステップと、

第1又は第2音声情報列の再生中は文字情報列

発声上の区切り又は言語上なんらかの区切りにもとづき、あるいはこれら区切りと特に係わりなく複数の区画に区分して記録するように働き、第2音声情報列を記録する第2の領域は、第1音声情報列の音声情報を例えば言い換えた音声情報を複数の区画に区分して記録するように働き、第3音声情報列を記録する第3の領域は、第1および第2音声情報列の複数の区画を単独又はひとまとめにして解説等をする音声情報を区画群で記録するように働き、記録位置識別情報を記録する第4の領域は、上記3つの情報列の媒体における記録位置を区画又は区画群ごとに識別させるように働く。

本発明に係る情報記録媒体の第3の態様は以上のように構成されるので、第1音声情報列を記録する第1の領域は、基本となる音声情報を例えば発声上の区切り又は言語上なんらかの区切りにもとづき、あるいはこれら区切りと特に係わりなく複数の区画に区分して記録するように働き、第2音声情報列を記録する第2の領域は、第1音声情報列の音声情報を例えば言い換えた音声情報を複

数の区画に区分して記録するように働き、文字情報列を記録する第3領域は、第1又は第2音声情報列に対応した文字情報を複数の区画に区分して記録するように働き、記録位置識別情報を記録する第4の領域は、上記3つの情報列の媒体における記録位置を区画ごとに識別させるように働く。

また、本発明に係る情報再生方式の第1の様態は以上のように構成されるので、第1および第2のステップは音声情報列の再生を指令された他の音声情報列の該当する区画のものに切り換えるよう働き、この切り換えに際して記録位置識別情報は切り換えられる情報列の記録位置を区画ごとに指示するように働く。

本発明に係る情報再生方式の第2の様態は以上のように構成されるので、第1、第2および第3のステップは音声情報列の再生を指令された他の音声情報列の該当する区画又は区画群のものに切り換えるよう働き、この切り換えに際して記録位置識別情報は切り換えられる情報列の記録位置を区画又は区画群ごとに指示するように働く。

2つの音声情報列を記録していることを特徴とする。すなわち、第1音声情報列は例えばネイティブスピーカーが自然な速さで話す英語の音声情報からなり、この音声情報列は複数の区画(いわばセグメント)に分割されている。第2音声情報列は上記第1音声情報列の内容に対応しているが別の音声情報であり、例えば単語を区切って話すゆっくりとした速さの英語の音声情報である。

ここで重要なことは、上記第1および第2音声情報列はそれぞれ複数の区画に区分されているが、互いに区画ごとに対応していることである。例えば、第1音声情報列のt番目の区画がネイティブスピーカーの話す"it's"であるときは、第2音声情報列のt番目の区画は単語を区切って話す

"it is"となる。従って、第1および第2音声情報列における区画とは、発声上の区切り又は言語上のなんらかの区切りにもとづいて分割したものであることを意味する。但し、これらの区切りにもとづいて区画が形成されることは、本発明に必須ではなく、後にD A Tを用いた例で示すよう

本発明に係る情報再生方式の第3の様態は以上のように構成されるので、第1および第2のステップは音声情報列の再生を指令された他の音声情報列の該当する区画のものに切り換えるよう働き、第4のステップは再生中の音声情報に対応する区画の文字情報をディスプレイ表示するように働き、このディスプレイ表示および再生の切り換えに際して記録位置識別情報はディスプレイ表示され又は再生が切り換えられる情報列の記録位置を区画ごとに指示するように働く。

[実施例]

以下、本発明の具体的な実施例を説明するに先立って、本発明の上記6つの様態のそれぞれについて、その基本的内容を説明する。なお、特許請求の範囲に記載された各構成要件が、これらによって限定されるものではないことは言うまでもない。

本発明に係る情報記録媒体の第1の様態は、特許請求の範囲第1項において構成要件が明らかにされている。この様態の記録媒体は、少なくとも

に、上記区切りと無関係な時間幅であってもよい。また、第2音声情報列と対応した内容でかつ別の音声情報からなるということは、言語上は同一の意味で発声の異なるものであることを示している、と言い換えることができる。

更にこの様態の記録媒体は、記録位置識別情報が記録されていることを特徴とする。そしてこの記録位置識別情報は、第1および第2音声情報列の音声情報が、媒体のどの位置に記録されているかを示す作用をもつ。従って、例えば第1音声情報列のt番目の区画の"it's"に対応する第2音声情報列の"it is"が、媒体のどの位置に記録されているかということは、この記録位置識別情報により認識することができる。

その結果、第1および第2音声情報列と記録位置識別情報は互いに無関係に記録されるのではなく、一定の関係をもって記録され、各情報は区画を単位として有機的に組み合わされていることがわかる。すなわち、第1および第2音声情報列は互いに対をなしており、これらを区画ごとに連関

させているのが記録位置識別情報である。なお、この態様の記録媒体は本発明の最も基本的な態様である。

本発明に係る情報記録媒体の第2の態様は、特許請求の範囲第7項において構成要件が明らかにされている。そして、この態様が上記第1の態様と異なる点は、第1および第2音声情報列の他に、第3音声情報列が媒体に記録されていることである。

ここで重要なことは、第3音声情報列は第1および第2音声情報列の1又は2以上の区画をひとまとまりとした区画群に区分されていることである。言い換えれば、第3音声情報列の1つの区画群は第1および第2音声情報列の1又は2以上の区画を包含しており、従って1つの区画群は1又は2以上の区画と対になっている。

更にこの態様の記録媒体では、記録位置識別情報は第3音声情報列の内容の記録位置を区画群ごとに示す作用をもっている。従って、第1、第2および第3音声情報列と記録位置識別情報は互い

に一定の関係をもって媒体に記録され、各情報は区画あるいは区画群を単位として有機的に組み合わせられることになる。

に一定の関係をもって媒体に記録され、各情報は区画あるいは区画群を単位として有機的に組み合わせられることになる。

本発明に係る情報記録媒体の第3の態様は、特許請求の範囲第13項において構成要件が明らかにされている。そして、この態様が上記第2の態様と異なる点は、第1および第2音声情報列の他に、文字情報列が媒体に記録されていることである。この文字情報列は第1又は第2音声情報列に対応する内容の文字情報からなり、例えばネイティブスピーカの話す英語を文字にて表わすものである。

この文字情報列についても、第1又は第2音声情報列の区画と対応する区画に区分されている。また、記録位置識別情報はこの文字情報列の記録位置を区画ごとに識別させるように機能する。従って、第1および第2音声情報列と文字情報列はそれぞれ区画単位で対応することになる。なお、この態様の記録方式において、第2の態様における第3音声情報列を記録情報として加えるときは、

第1および第2音声情報列と文字情報列の1又は2以上の区画は第3音声情報列の1つの区画群に対応することになる。

以上が、本発明に係る情報記録媒体の3つの態様である。これに対して、本発明に係る情報再生方式には次の3つの態様があり、これらの態様はそれぞれ先に示した情報記録媒体の3つの態様に對比される関係にある。

本発明に係る情報再生方式の第1の態様は、特許請求の範囲第19項において構成要件が示されている。この態様の再生方式では、記録媒体の第1の態様に係る音声情報が媒体にあらかじめ記録されていることを前提とする。そして更に、下記の2つのステップを少なくとも備えることを特徴とする。

第1のステップとは、第1の音声情報列から第2の音声情報列への再生の切り換えである。そして、この切り換えは区画を単位としてなされる。例えば、第1音声情報列のt番目の区画が再生されているときに第2音声情報列の再生指令が入力

されると、記録位置識別情報にもとづき第2音声情報列のt番目の区画を読み出し、その音声情報を再生する。第2のステップとは、第2音声情報列から第1音声情報列への再生の切り換えであり、その切り換えも第1のステップと同様に区画に従ってなされる。

なお、第1および第2のステップは上記のものに限られず、種々の変形が可能である。その代表的なものとして、いわゆる戻し指令がある。すなわち、再生中の停止命令により一時再生を中断した後に戻し指令が入力されたときは、指令された直だけ音声情報の読み出し位置を戻すことにより、より希望に合った音声情報の再生を行うものである。

本発明に係る情報再生方式の第2の態様は、特許請求の範囲第22項において構成要件が示されている。そして、この態様が上記第1の態様と異なる点は、記録媒体の第2の態様に係る音声情報が媒体にあらかじめ記録されていることを前提とすることと、第1又は第2音声情報列から第3音

声情報列への再生の切り換えをする第3のステップを備えていることである。

例えば、第1音声情報列の再生中にネイティブスピーカの“it's”が聴き取れなかったときは、第1のステップによって第2音声情報列に再生を切り換えれば、ゆっくりと単語を区切って話す“it's”を聴くことができる。そして、この日本語の意味や文法を知りたいときは、第3のステップによって第3音声情報列に再生を切り換えればよい。

もちろん、この態様の方式においても、上記第1の態様で説明した戻し指令や停止命令を組み合わせて使えるよう変形できることは言うまでもない。

本発明に係る情報再生方式の第3の態様は、特許請求の範囲第26項において構成要件が示されている。そして、この態様が上記第2の態様と異なる点は、記録媒体の第3の態様に係る音声情報、文字情報が媒体にあらかじめ記録されていることを前提とすることと、第1又は第2音声情報列の再生中に文字情報列のディスプレイ表示がなされ

るようになっていていることである。

例えば、第1音声情報列の“it's”が再生されているときは、所定の表示部に“it's”もしくは“it's”がディスプレイ表示される。なお、この表示については再生中の音声情報と時間的に完全に同期している必要はなく、文字が少しずつ遅れて表示されたり、あるいは少しずつ先に表示されたりしてもよい。

次に、添付図面を参照して本発明のいくつかの実施例を具体的に説明する。

まず最初に、第1図ないし第3図を参照して本発明に係る記録媒体の第1の例を説明する。第1図は本発明の記録媒体を英会話独習用に適用したときの情報列A、B、Cとその記録内容の説明図である。同図において、情報列Aはネイティブスピーカの話す英語の音声情報列（第1音声情報列）を示し、複数に区分されたセグメント621～627はその区画を示す。情報列Bはゆっくりと単語を区切って話す英語の音声情報列（第2音声情報列）を示し、これはセグメント621～

627に対応する英単語、句からなる区画の音声情報である。情報列Cは日本語の解説をする音声情報列（第3音声情報列）を示し、情報列A、Bのセグメント621～624、625～627にそれぞれ対応している。

第2図は第1図に示す例のセグメント当りの時間と容量の関係の説明図である。同図において、1秒間は6キロバイトの容量に対応している。例えばセグメント623の“much of a”はネイティブスピーカの話す速さで0.4秒かかり、従って媒体で占める記録容量は2.4キロバイトである。

第3図は第1図および第2図に示す例のディレクトリの一例の説明図である。同図において、ディレクトリは1セグメントあたり $9 \times 3 = 27$ バイトで構成される。情報列A、B、Cはそれぞれ第1図の情報列A、B、Cに対応している。また、1バイトのCは属性を示し、これが129のときすなわちビット表現（8ビット）で“10000001”のときは前の情報列の内容と同じ内容で

あることを示す。

位置情報のM、S、B（各1バイト）は産業界で標準になっているCD-ROM上の位置を表わすパラメータである。すなわちMは分、Sは秒、Bはブロックをそれぞれ示し、1ブロックは2048バイトからできている。そして75ブロックで1秒分を構成している。従って、最大の数値はM=59、S=59、B=74となる。次の2バイトのSBはスタートバイトを示し、その次の3バイトのLLLはバイト長を示す。なお、位置を示すパラメータに分、秒を使う理由はCD-ROMはもともと音楽用として開発されたためであり、従って始めからの時間として記録位置を表現するようになっていている。そのためCD-ROMとした場合には、この分と秒は再生時の時間とは全く無関係であり、単に位置を表わしている情報にすぎないことになる。

その結果、例えば623セグメントの“much of a”は、0分11秒42ブロックの1354バイト目から2400バイトの長さでネイティブス

ピーカの話す英語が記録され、0分11秒43ブロックの1706バイト目から7200バイトの長さでゆっくりと単語を区切って話す英語が記録され、0分11秒6ブロックの1282バイト目から72000バイトの長さで日本語解説が記録される。なお、621~627のセグメントナンバーはメモリ上ではなく、そのアドレスに対応している。

次に、上記の第1図ないし第3図に示す記録媒体の作用を説明する。

第3図に示すように、媒体の0分11秒3ブロックにおける826バイト目から826+1200-1=2025バイト目までの領域には、セグメントが621で属性Cが0の情報列すなわちネイティブスピーカが話す "It's" を記録するように働く。また、媒体の0分11秒3ブロックにおける2026バイト目から2026+5400-1=7425バイト目までの領域すなわち0分11秒6ブロックの1282バイト目までの領域には、セグメントが621で属性Cが

64の情報列すなわち単語を区切ってゆっくり話す "It's" を記録するように働く。また、媒体の0分11秒6ブロックにおける1282バイト目から0分11秒41ブロックにおける1601バイト目までの領域には、セグメントが621で属性Cが128の情報列すなわち日本語の解説を記録するように働く。

このように、第3図に示すようなディレクトリを形成すれば、第2図に示すような容量および再生時間で第1図に示す情報列を記録できる。

次に、第4図ないし第9図を参照して、第1図ないし第3図により説明した記録媒体からの再生方式を説明する。

第4図は再生装置の一例の構成を示すブロック図である。同図において、媒体1は例えばCD-ROMであり、再生機構2にセットされる。再生機構2はディスクインターフェイス(I/F)3およびバス4を介してCPU5に接続される。また、バス4にはプログラムを格納するための例えば32キロバイトのROM6と、ディレクトリや

音声情報列を一時的に格納するための例えば256キロバイトのRAM7とが接続されている。さらに、バス4には手動操作のためのハンドセット8との間で情報の授受を行なうハンドセットインターフェイス(I/F)9と、AMP10を介して外部端子11およびハンドセット8に接続されたD/Aコンバータ12に接続されている。なお、ハンドセット8にはイヤホン13が接続されている。

第5図(a)、(b)はROM6およびRAM7のメモリ割り当てを説明する図である。同図(a)に示す如く、ROM6の32キロバイトにはプログラムが格納される。同図(b)に示す如く、RAM7は(50+50)=100キロバイト分の50ブロック用バッファと、(75+75)=150キロバイト分のディレクトリと、6キロバイト分のシステムエリアに割り当てられる。従って、RAM7には常時50ブロック分の音声情報列が保持され、かつ150キロバイト+27バイト=5555セグメント分のディレクトリ(情

報列Aの部分のみで約30分間に相当)が保持される。

なお、上記の例では媒体としてCD-ROMを用いているが、その代表的なものの容量は552メガバイトである。CD-ROMではアドレスを表わすのに分、秒、ブロックの単位を用いている。前述したように1ブロックが2048バイト、75ブロックが1秒、60秒が1分で構成され、最大59分59秒74ブロックである。従って

$$= 552.96 \text{ メガバイト}$$

となる。このうち、最初から2秒分はCD-ROMのフォーマットとしてユーザは使えないので、正確には最大容量とし552.6528MBとなる。このうち、最初から20秒に相当するところにディレクトリを割り当てると、3メガバイトのディレクトリ容量をCD-ROMに確保することができる。

次に容量に関する計算例を示す。

音のサンプルをADPCM方式によって16キ

ロサンプル/秒とし、1サンプル当り3ビットとする。このようにすると48キロビット/秒であり、従って6キロバイト/秒となり、転送レートもこれに合わせる必要がある。なお、16キロサンプル/秒であれば8KHzまでのf特性がある。従って、子音まで十分に記録できる。

このように仮定すると、1時間の録音に6キロバイト×3600秒=21.6メガバイトの容量が必要となる。一般に、CD-ROM1枚にはエラーコレクションを入れて552メガバイト記録できる。ディレクトリ部を除くと549メガバイトを音声情報の収納部として使える。従って、 $549 + 21.6 = 570.6$ 時間24分の音声情報を記録することが可能となる。そこで、英会話用として用いる場合には、ネイティブスピーカによる自然の速度で発音されている物語を1時間とすると、それを単語1つ1つを区切ってゆっくり発音している部分はその4倍で4時間とみることができる。そして、解説の部分は合計で15時間あったとしても全部で20時間である。

第7図(a)は上記再生装置の斜視図であり、第7図(b)はこれをベルト28付のケース29に収納したときの斜視図である。CD-ROMは近年の技術進歩により小型、軽量化しているので、図示の如く携帯型にできる。第8図はハンドセットを手を持った状態を示している。

次に、上記再生方式における再生シーケンスの一例に従って、本実施例の作用を説明する。第9図はそのシーケン図である。情報列Aの部分すなわちネイティブスピーカの英語を続けて聴くときは、第9図(a)のようになる。このときは、第6図のボタン23Aを押してそのまま聴いているだけでよい。

次に、第9図(b)のシーケンスに従って説明する。まず、セグメント621から情報列Aの再生を行う。ここで、セグメント623の“much of a”がよく聴き取れなかった時には、すぐにSTOPボタン24を押す。このとき、セグメント表示部21のセグメント番号は624になっている。そこで、REVボタン25を1回だけたたく。

次に、1時間の会話部分をいくつかの区画(セグメント)に分解できるかを考えると、1秒間に平均4個の区切りとして1時間で14400個である。1個の区切りのディレクトリには27バイト必要なので全体で約389キロバイトであり、これは上記3メガバイトのディレクトリの収納場所に充分収納でき、従って1時間の物語の全てのディレクトリを格納することができる。

第6図は第4図に示すハンドセット8とイヤホン13の詳細な説明図である。ハンドセット8の前面にはセグメント番号を表示するセグメント表示部21と、早送りのためのFFスイッチ22と、それぞれ音声情報列A、B、Cの再生指令をするためのA、B、C指令スイッチ23A、23B、23Cと、再生の停止を指令するためのSTOPボタン24と、再生の戻しを指令するためのREVボタン25とを備えている。また、イヤホン13はコード26を介してハンドセット8に接続され、ハンドセット8はコード27を介して本体に接続されている。

なお、REVボタン25は押しつづけていると連続してもどり、1回たたき毎に表示部21のセグメントは1つだけ戻る。セグメントが624の状態ではREVボタン25を1回だけたたくと、表示部21のセグメントは623になる。

次に、スイッチ23Bを押すと“much/of/a”と1つ1つの単語を区切ってゆっくり発音しているのが聴える。そのままにしておくと“problem”と続けて情報列Bのまま進む。“much/of/a”を聴いたところでSTOPボタン24を押し、再度REVボタン25を押してセグメントを623にする。そしてスイッチ23Cを押すと、情報列Cの日本語の解説を聴くことができる。

解説のところはいくつかの単語がまとまった句について解説していることもあるので、いくつかのセグメント番号で一つの解説ということもある。属性Cを示すバイトの最下位ビットが1のものは(属性Cが129のもの)はそれを示し、同一の情報列の前のセグメント番号と同じ内容であることを意味している。ここに当たったら読むとばすこ

ともできる。

次に、第10図ないし第12図を参照して本発明に係る記録媒体と再生方式の第2の例を説明する。第10図はディレクトリの説明図で、1セグメント当り12バイトで構成される。すなわち、1バイトの属性Cと、2バイトのセグメントナンバーSSと、1バイトのサフィックスナンバーNと、1バイトの分Mと、1バイトの秒Sと、1バイトのブロックBと、2バイトのスタートバイトSBと、3バイトのバイト長LLLの計12バイトである。

また、属性Cの8ビットについて、第1ビット（最上位ビット）が1のときはセグメントのスタートを示し、0のときはその他の状態であることを示す。第2、3ビットが0のときは情報列Aであることを示し、1のときは情報列Bであることを示し、2のときは情報列Cであることを示す。第4ビットが0のときはサフィックスが無いことを示し、1のときはサフィックスが有ることを示す。第5ビットが1のときは前のセグメントと同

じ内容であることを示し、0のときは同じ内容でないことを示す。

第11図は第10図のディレクトリに従った情報列A、B、Cのセグメント分割を示す。そしてこれが前述の第1の実施例と異なる点は、サフィックスを付したセグメントで分割していることである。第12図は8ビットの属性Cの第2、3ビットの値とセグメントの関係図である。

次に、第13図および第14図を参照して本発明の第3の例に係る記録媒体と再生方式を説明する。そしてこの例が第1および第2の例と異なる点は、音声情報だけでなく文字情報の記録、再生をも行なうようにしていることである。

第13図はそのディレクトリの説明図である。そしてこれが第10図のものと異なる点は、属性Cの第2、3のビットにD=3の情報が付加されることである。このビットが3であるときは、文字情報列Dが媒体にコードで記録されていることを示す。

第14図はこの例に係る再生装置に用いられる

ハンドセットの正面図である。そしてこれが第6図のものと異なる点は、例えばLCDで構成される文字表示部41と、文字の表示をオン、オフさせるための文字表示ボタン42が設けられていることである。このようなハンドセット8において文字表示ボタン42を押すと、文字表示部41で文字情報列Dが表示されたり表示されなかったりする。

次に、この第3の実施例の作用の特色を説明する。

音声情報列Aの再生中は情報列Aの発音のスピード、すなわちLLLの長さに応じてその文字情報列Dの現われるスピードも制御する。つまり、このセグメントの発音の始まりから終わりまでの間に、文字が表示部41に現われ始めて完全に出力し終るようにする。すなわち、発音に完全に同期させて文字を表示させる。

次に、音声情報列Bの再生中は音声情報列Aの部分の再生に比べて数倍の時間をかけて発音されているので、この時の情報列Bの発音の長さに同

期させて文字情報列Dを表示させる。実際には、文字の表示の方を早めに出す（音声出力の方を少し遅らせて出す）方が使う人にとって便利なことも考えられるが、このことについては実用化に際して更に検討してみる必要がある。

次に、本発明の第4の実施例として、情報記録媒体をランダムアクセスのできない磁気テープとした場合を説明する。

この例においても、基本的な情報の記録、再生の様子は前述の例と同様であるが、磁気テープではデータの飛び越し再生が容易でないため、実用的な再生装置とするためには、情報を一時的に記憶しておくためのバッファを設けるのが望ましい。具体的には、前述の仮想トラックの情報列A、B、C（3トラック）に相当する各小単位を、 A_1 、 B_1 、 C_1 、 A_2 、 B_2 、 C_2 、 A_3 、 B_3 、 C_3 、…というように、対応する情報列A、B、Cを1つのかたまりとして順番にならべる。そして、情報列Aを再生する時は A_1 、 A_2 、 A_3 、…というようにAのみ取り出して再生し、情報列

Bを再生する時は B_1 、 B_2 、 B_3 、…というようにBのみ取り出して再生する。情報列Cについても同様である。この時、 A_1 の次に A_2 を切れ目なく再生するためには、 B_1 、 C_1 を飛びこさなければならない。これは、通常のテープ走行では容易でないため、磁気テープからは早めにバッファの中に情報を取ってきておく。そして、飛びこして再生するのはバッファの中で行うようにする。

この磁気テープを、いわゆるDAT（デジタル・オーディオ・テープ）としたときには、次のようになる。まず、区画として、Aトラックに記録されるネイティブ・スピーカの発音をある単位で区切る。例えば、ここでは1秒毎の一定時間に区切ることにする。そして、トラックBとしては、ゆっくり（例えば平均 $1/3$ の速度で）話している内容のものとする。Aトラックの1秒に相当するところをBトラック用に4秒分とれるよう領域を確保する。そして、解説用のCトラックとしては9.4秒分とれるよう領域を確保する。なお、

$86,400$ （バイト $+6000$ ）バイト）
 $=14.4$ sec
 を記録できる。

次に、バッファについての考察をする。

1MBのRAMを使用すると

$1,0480,576+86,400$
 $=12,135$

となるので、12個のバッファを設けることができる。12個のバッファを機能的な意味でリング状にならべ、常に2個づつ先取りしてテープからバッファにうつす。再生はバッファから取り出して行う。任意のところで止めた時には、10個のバッファが止めた時点より前の音声（既に再生音としてスピーカから出してしまったもの）が残っている。すなわち、Aトラック上の音声にして10秒の音声が入ったテープを巻き戻すことなく、バッファから繰返し再生することができる。

このようにすることにより、10個のバッファに残っている10秒間の音声情報については、前述したCD-ROMを用いた実施例と同じく、任

上記の秒数は一例であり、これに限定されるものではない。つまり、14.4秒を1つの単位とする。すなわち、14.4秒はA、B、Cトラックの各対応する3個のセグメントをひとまとめにしたものといえる。ここで注意を必要とするは、Aトラック用の1秒分はオリジナルのものを1秒づつ切っているだけで、もとの1秒と完全に一致していることである。Bトラックの4秒は4秒以内であればよく、3秒で終わっていて1秒あまってもよい。なお、この余ったところ（1秒分）を再生する時に飛ばす方法については後述する。

ロータリ・ヘッド型DATを例にとれば、次のようになる。この例では、ロータリ・ヘッドの1回転で2トラックを記録又は再生する方式が一般的である。そして、1トラックに2880バイトの音声情報を記録できるようになっている。従って、30トラックを1区切りとすると、86,400バイトとなる。この中に、前述の実施例でも使用した音声のサンプリング・レート $=48$ (Kbit/sec) で音声を記録すると、

意のセグメントにおいてA、B、Cのどの仮想トラックにも容易に移ることができる。

ただし、CD-ROMのようにランダム・アクセスができないので、A、B、Cの3つのトラックの対応するセグメントの時間数が、常に14.4秒以内にしなければならない。この時、仮想トラックAの部分は、常に一定の時間幅（例えば1秒毎）になるように、機械的に区切られていてもよい。例えば、Bトラックの部分は2秒の時もあるし5秒の時もあるので、可変になるようにする。14.4秒からA、Bトラックの時間を引いた時間がCトラックに割り当てられる時間である。なお、B、Cトラックのいずれも、区切りの中のAトラックに完全に対応した部分を記録する必要はなく、近くの複数の区切りを通して、あるまとまりができていればよい。

DATにおいてはサブコード領域があるので、各区切り（14.4秒/30トラック毎の）と、各区切りの中のA、B、Cトラックの境目を示すバイト数の情報は、このサブコード領域に記録す

ることができる。

次に、D A Tにどれだけの音声情報が入るかを具体的に計算する。

一般に、1時間用のD A Tには240,000トラックが記録できる。30トラックを1区切りに行っているが、次の6トラックを休みのトラックに使い、テープを止めている時にロータリーヘッドが多数回過っても、音声情報の記録されている部分にキズをつけないようにする。このようにするためには、36トラックが一単位となる。つまり、全体では6,666単位が記録でき、 $6,666 \times 14.4 \text{ 秒} = 26 \text{ 時間 } 39 \text{ 分 } 50 \text{ 秒}$ となる。このうち、Aトラック分として1時間に割り当てるので、B, Cトラックに25時間39分50秒分50秒を使えることになる。ゆっくり発音したBトラックが4時間分使ったとしても、説明用のCトラックに21時間以上割り当てることができる。充分な長さがある。

なお、上記のA, B, Cトラックのセグメントのとり方は、前述のC D・R O Mを応用したもの

面と同期させればよい。

また、パソコンなどにおいて本発明を用いることもできる。すなわち、C A I (Computer Aided Instruction) との連動やC D I (Compact Disk Interactive) 等高度のプログラマブルのものとも結合すれば、より高度の応用が可能である。また、パソコンの記憶装置を上記のC DやD A Tの代りに用いて本発明を実施することも可能である。

さらに、英会話などに限らず、詩吟、法律用のものにも適用することができる。さらにまた、情報列は3種又は4種のものに限らず、2種あるいは5種以上としてもよい。

〔発明の効果〕

以上、詳細に説明したように本発明に係る情報記録媒体によれば、上級者にも初級者にも本人の望む適切な内容の音声情報等を記録することができ、従ってこれらの者を共に満足させる情報記録媒体を得ることができる。

また、本発明に係る情報再生方式によれば、本発明に係る記録媒体を用いて、上級者にも初級者

と同じくAトラックは1秒等一定時間幅で区切るだけでなく、1秒程度を目安とし、何らかの発音上の切り目があるところで切ってもよい。すなわち、記録する時の編集方針により変化してもよい。従って、C D・R O Mのようにランダム・アクセスのできるものでも、一定時間幅で区切って区画を作ってもよいことになる。

本発明は上記実施例に限定されるものではなく、種々変形が可能である。

例えばビデオ・ディスクやビデオ・テープとの連動でも本発明を適用できる。すなわち、普通のスピード(ネイティブ・スピーカのスピード)のAトラックに例えば映画のサウンド・トラック部を入れておき、Bトラックにゆっくりした発音を入れておき、Cトラックに日本語解説を入れておく。そして、Aトラックの再生を行ないながら、理解できなかったところで止め、少しもどし(この時、画面は固定のままでも実用に耐える)でBトラックを聴き、それでも分らなければCトラックを聴く。なお、Aトラックを再生する時だけ画

にも適切な内容の音声情報等を再生して与えることができ、従ってこれらの者を共に満足させる情報再生方式を得ることができる。

4. 図面の簡単な説明

第1図は本発明の記録媒体を英会話独習用に適用したときの情報列A, B, Cとその記録内容の説明図、第2図は第1図に示す例のセグメント当りの時間と容量の関係の説明図、第3図は第1図および第2図に示す例のディレクトリの一例の説明図、第4図は再生装置の一例の構成を示すブロック図、第5図(a)、(b)はR O M 6およびR A M 7のメモリ割り当てを説明する図、第6図は第4図に示すハンドセット8とイヤホン13の説明図、第7図は第4図に示す再生装置の斜視図、第8図はハンドセットを手に持った状態を示す図、第9図は作用を説明するシーケンス図、第10図は本発明の第2の例に係るディレクトリの説明図、第11図は第10図のディレクトリに従った情報列A, B, Cのセグメント分割を示す図、第12

図は8ビットの属性Cの第2, 3ビットの値とセグメントの関係図、第13図は本発明の第3の例に係るディレクトリの説明図、第14図はこの例に係る再生装置に用いられるハンドセットの正面図である。

1…媒体、2…再生機構、8…ハンドセット、
13…イヤホン、21…セグメント表示部、
22…FFスイッチ、23A, 23B, 23C…
A, B, C指令スイッチ、24…STOPボタン、
25…REVボタン、41…文字表示部、42…
文字表示ボタン。

特許出願人 カナース・データー株式会社
代理人弁理士 長谷川 芳 樹

	記 録 内 容
A	<p>" It's not much of a problem. I'd second that. "</p> <p>621 622 623 624 625 626 627</p>
B	<p>" It is not much of a problem. I would second that. "</p> <p>621 622 623 624 625 626 627</p>
C	<p>(621~624)</p> <p>「"It"とは前の〇〇を指しており、"not much of a problem"は慣用句で「問題がない」とか「心配ない」というような意味です。</p> <p>(625~627)</p> <p>次に、"I'd"とは"I would"をつめた言い方であり、同じように"I could"も"I'd"とつめて言えます。 "second"とは「支持する」、「賛成する」という意味があり、"I'd second that"で「同感です」という意味になります。」</p>

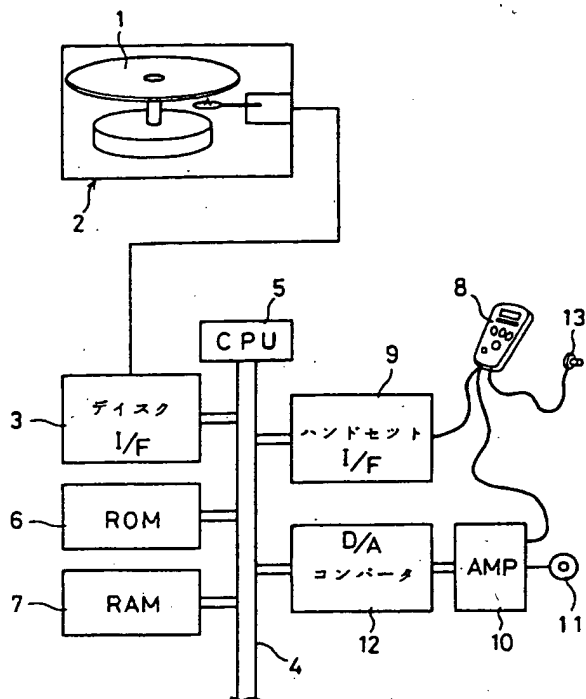
第 1 図

		621	622	623	624	625	626	627	計
時間 (秒)	A	0.2	0.1	0.4	0.3	0.1	0.3	0.4	1.8
	B	0.9	0.2	1.2	0.6	0.8	0.6	0.7	5.0
	C	12				25			37
容量 (KB)	A	1.2	0.6	2.4	1.8	0.6	1.8	2.4	10.8
	B	5.4	1.2	7.2	3.6	4.8	3.6	4.2	30.0
	C	72				150			222

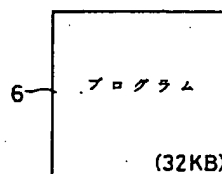
第 2 図

情報列	セグメント	カテゴリー	位置情報			スタートバイト	長さ (BYTE)
	SS	C	M	S	B	SB	LLL
A	621	0	0	11	3	826	1,200
B		64	0	11	3	2,026	5,400
C		128	0	11	6	1,282	72,000
A	622	0	0	11	41	1,602	600
B		64	0	11	42	154	1,200
C		129	0	11	6	1,282	72,000
A	623	0	0	11	42	1,354	2,400
B		64	0	11	43	1,706	7,200
C		129	0	11	6	1,282	72,000
A	624	0	0	11	47	714	1,800
B		64	0	11	48	466	3,600
C		129	0	11	6	1,282	72,000
A	625	0	0	11	49	2,018	600
B		64	0	11	50	570	4,800
C		128	0	11	52	1,274	150,000
A	626	0	0	12	50	1,770	1,800
B		64	0	12	51	1,522	3,600
C		129	0	11	52	1,274	150,000
A	627	0	0	12	53	1,026	2,400
B		64	0	12	54	1,378	4,200
C		129	0	11	52	1,274	150,000

第 3 図



第 4 図

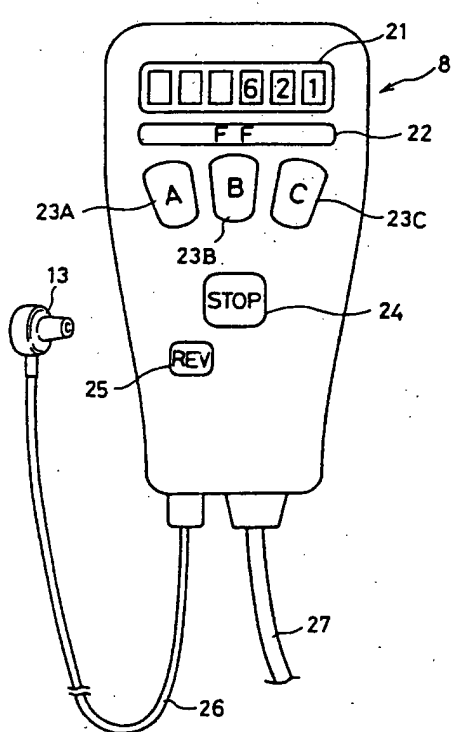


(a)

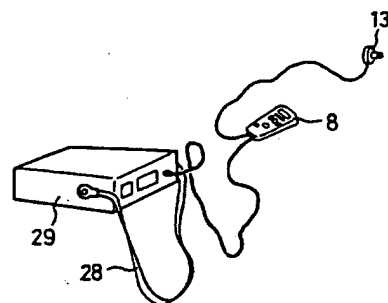
トラック用	(50KB)
ダブルバッファ	(50KB)
ディレクトリ用	(75KB)
ディレクトリ用	(75KB)
システムエリア	(6KB)

(b)

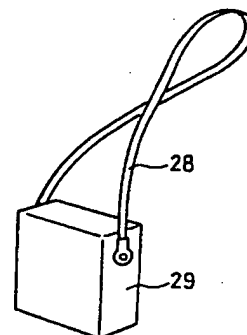
第 5 図



第 6 図

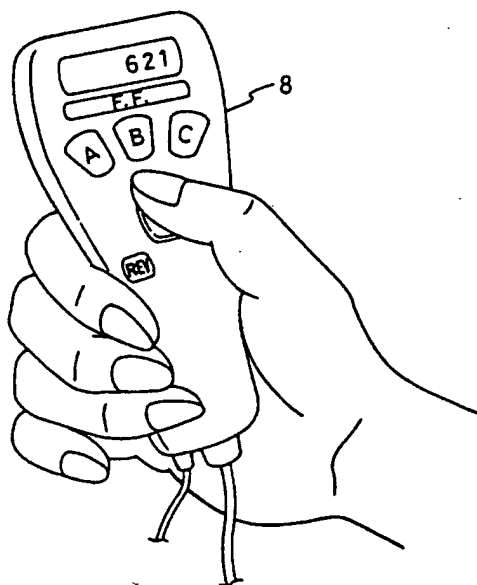


(a)

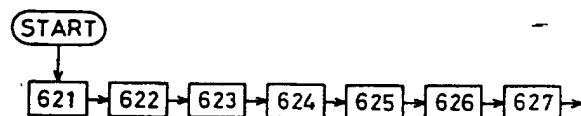


(b)

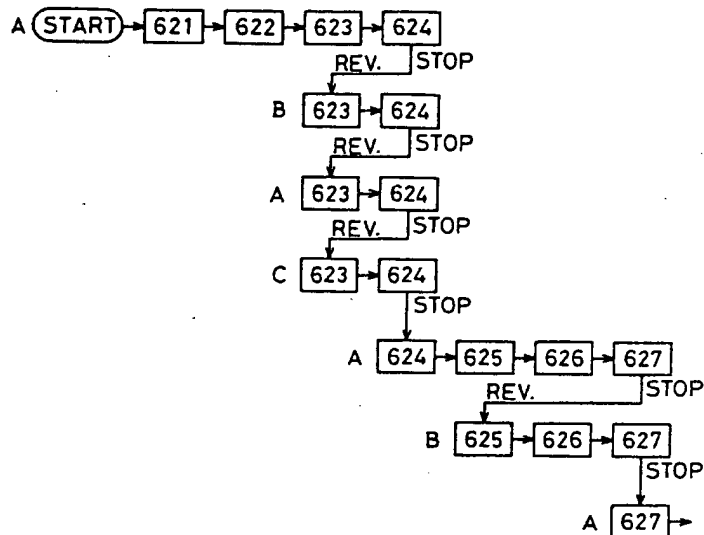
第 7 図



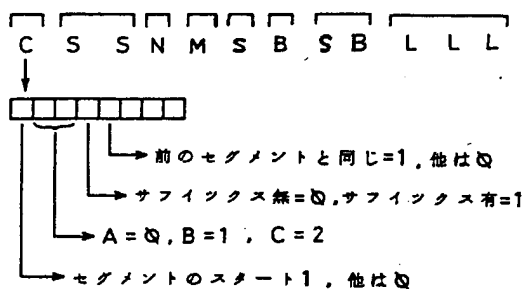
第 8 図



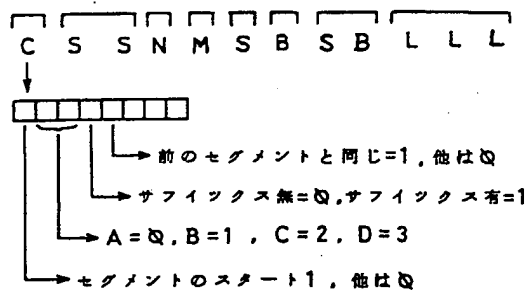
第 9 図 (a)



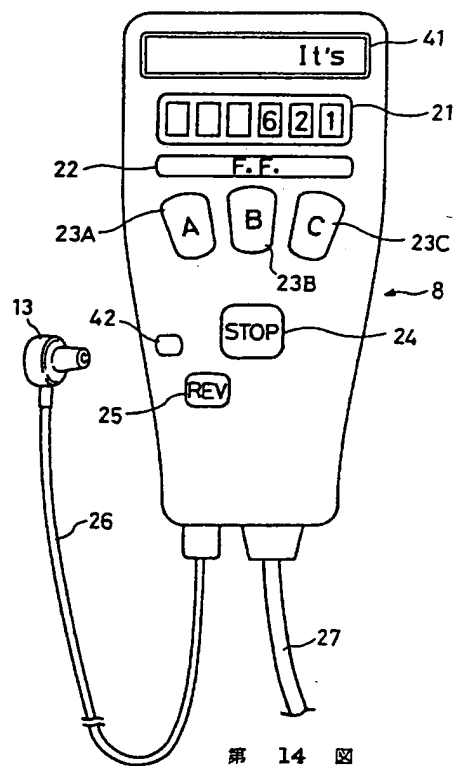
第 9 図 (b)



第 10 図



第 13 図



第 14 図

情報列	セグメント分割
A	<p>It's not much of a problem. I'd second that.</p> <p>621 622 623 624 625 626 627</p>
B	<p>It is not much of a problem. I would second that.</p> <p>621-1 621-2 622 623-1 623-2 623-3 624 625-1 625-2 626 627</p>
C	<p>621-1 621-2 622 623-1 623-2 623-3 624 625-1 625-2 626 627</p> <p>□ □ □ □ □ □ □ □ □ □ □</p>

第 11 図

2, 3ビット目	セグメント
0	621
1	621-1
1	621-2
2	621-1
3	621-2
0	622
1	622
2	622
0	623
1	623-1
1	623-2
1	623-3
2	623-1
2	623-2

第 12 図

Date: December 20, 2003

Declaration

I, Michihiko Matsuba, President of Fukuyama Sangyo Honyaku Center, Ltd., of 16-3, 2-chome, Nogami-cho, Fukuyama, Japan, do solemnly and sincerely declare that I understand well both the Japanese and English languages and that the attached document in English is a full and faithful translation, of the copy of Japanese Unexamined Patent No. Sho-63-191372 laid open on August 8, 1988.

A handwritten signature in black ink, appearing to read "m. matsuba". The signature is fluid and cursive, with the first letter of "matsuba" being a large, stylized "m".

Michihiko Matsuba

Fukuyama Sangyo Honyaku Center, Ltd.

DATA RECORDING MEDIUM AND DATA REPRODUCTION SYSTEM

Japanese Unexamined Patent No. Sho-63-191372

Laid-open on: August 8, 1988

Application No. Sho-62-219962

Filed on: September 2, 1987

Inventor: Hiroshi SEKIGUCHI

Applicant: Kanars Data Corporation

Patent Attorney: Yoshiki HASEGAWA et al.

SPECIFICATION

1. TITLE OF THE INVENTION

Data Recording Medium and Data Reproduction System

2. WHAT IS CLAIMED IS;

1. A data recording medium for recording speech sound data, at least comprising:

a first region for recording a first speech sound data sequence divided into a plurality of sections;

a second region for recording a second speech sound data sequence which has contents corresponding to said first speech sound data sequence, which is made up of different speech sound data and which is divided into a plurality of sections

corresponding to the sections of said first speech sound data sequence; and

a third region for recording recorded position identification data which indicates the recorded positions of said first and second speech sound data sequences, respectively, using each of said sections.

2. The data recording medium according to Claim 1, wherein said third region is a region for collectively recording at least a portion of said recorded position identification data at a predetermined position as a directory.

3. The data recording medium according to Claim 1, wherein said third region is a region for recording, as identification data, at least portions of said recorded position identification data at positions adjacent to the respective recorded positions of said first and second speech sound data sequences.

4. The data recording medium according to Claim 1, wherein each of said sections in said first and second regions is made up of one, or more, small sections, and

said third region includes recorded position identification data for each of said small sections.

5. The data recording medium according to Claim 1, wherein said data recording medium is a disk-type recording medium such

as a CD.

6. The data recording medium according to Claim 1, wherein said data recording medium is a tape-type recording medium such as a DAT.

7. A data recording medium for recording speech sound data, at least comprising:

- a first region for recording a first speech sound data sequence divided into a plurality of sections;

- a second region for recording a second speech sound data sequence which has contents corresponding to said first speech sound data sequence, which is made up of different speech sound data and which is divided into a plurality of sections corresponding to the sections of said first speech sound data sequence;

- a third region for recording a third speech sound data sequence which has contents corresponding to said first and second speech sound data sequences, which is made up of different speech sound data and which is divided into a plurality of section groups, each group made up of one, or more, sections from among said sections; and

- a fourth region for recording recorded position identification data which indicates the recorded positions of said first and second speech sound data sequences, respectively,

using each of said sections and for recording recorded position identification data which indicates the recorded positions of said third speech sound data sequence using each of said section groups.

8. The data recording medium according to Claim 7, wherein said fourth region is a region for collectively recording at least a portion of said recorded position identification data at a predetermined position as a directory.

9. The data recording medium according to Claim 7, wherein said fourth region is a region for recording, as identification data, at least portions of said recorded position identification data at positions adjacent to the respective recorded positions of said first, second and third speech sound data sequences.

10. The data recording medium according to Claim 7, wherein each of said sections in said first and second regions is made up of one, or more, small sections, and

said fourth region records recorded position identification data for each of said small sections.

11. The data recording medium according to Claim 7, wherein said data recording medium is a disk-type recording medium such as a CD.

12. The data recording medium according to Claim 7, wherein

said data recording medium is a tape-type recording medium such as a DAT.

13. A data recording medium for recording speech sound data and text data, at least comprising:

a first region for recording a first speech sound data sequence divided into a plurality of sections;

a second region for recording a second speech sound data sequence which has contents corresponding to said first speech sound data sequence, which is made up of different speech sound data and which is divided into a plurality of sections corresponding to the sections of said first speech sound data sequence;

a third region for recording a text data sequence which is made up of text data having contents corresponding to said first and second speech sound data sequences and which is divided into a plurality of sections corresponding to the sections of said first speech sound data sequence; and

a fourth region for recording recorded position identification data which indicates the respective recorded positions of said first and second speech sound data sequences and a text data sequence using each of said sections.

14. The data recording medium according to Claim 13, wherein said fourth region is a region for collectively recording at

least a portion of said recorded position identification data at a predetermined position as a directory.

15. The data recording medium according to Claim 13, wherein said fourth region is a region for recording, as identification data, at least portions of said recorded position identification data at positions adjacent to the respective recorded positions of said first and second speech sound data sequences as well as a text data sequence.

16. The data recording medium according to Claim 13, wherein each of said sections in said first, second and third regions is made up of one, or more, small sections, and

said fourth region is a region for recording recorded position identification data for each of said small sections.

17. The data recording medium according to Claim 13, wherein said data recording medium is a disk-type recording medium such as a CD.

18. The data recording medium according to Claim 13, wherein said data recording medium is a tape-type recording medium such as a DAT.

19. A data reproduction system for reproducing speech sound data that has been recorded on a medium in advance, wherein said speech sound data at least includes a first speech sound data sequence divided into a plurality of sections; a second

speech sound data sequence which has contents corresponding to said first speech sound data sequence, which is made up of different speech sound data and which is divided into a plurality of sections corresponding to the sections of said first speech sound data sequences; and recorded position identification data which indicates recorded positions of said first and second speech sound data sequences, respectively, in said medium using each of said sections, and

the data reproduction system at least comprises:

the first step of reading out speech sound data of a section in said second speech sound data sequence corresponding to the section in said first speech sound data sequence during reproduction from said medium based on said recorded position identification data when a reproduction instruction of said second speech sound data sequence is inputted during the reproduction, or after an interruption, of said first speech sound data sequence and of reproducing the read out speech sound data; and

the second step of reading out speech sound data of a section in said first speech sound data sequence corresponding to the section in said second speech sound data sequence during reproduction from said medium based on said recorded position identification data when a reproduction instruction of said

first speech sound data sequence is inputted during the reproduction, or after an interruption, of said second speech sound data sequence and of reproducing the read out speech sound data.

20. The data reproduction system according to Claim 19, wherein said first step is the step of shifting back the position of the speech sound data in said second speech sound data sequence that has been read out from said medium by an instructed amount in units of a notch with respect to said sections when a return instruction is inputted at the time of input of said reproduction instruction.

21. The data reproduction system according to Claim 19, wherein said second step is the step of shifting back the position of the speech sound data in said first speech sound data sequence that has been read out from said medium by an instructed amount in units of a notch with respect to said sections when a return instruction is inputted at the time of input of said reproduction instruction.

22. A data reproduction system for reproducing speech sound data that has been recorded on a medium in advance, wherein said speech sound data at least includes: a first speech sound data sequence divided into a plurality of sections; a second speech sound data sequence which has contents

corresponding to said first speech sound data sequence, which is made up of different speech sound data and which is divided into a plurality of sections corresponding to the sections of said first speech sound data sequence; a third speech sound data sequence which has contents corresponding to said first and second speech sound data sequences, which is made up of different speech sound data and which is divided into a plurality of section groups, each group made up of one, or more, sections from among said sections; and recorded position identification data which indicates the recorded positions of said first and second speech sound data sequences, respectively, in said medium using each of said sections and which indicates the recorded position of said third speech sound data sequence in said medium using each of said section groups, and

the data reproduction system at least comprises:

the first step of reading out speech sound data of a section in said second speech sound data sequence corresponding to the section in said first speech sound data sequence during reproduction from said medium based on said recorded position identification data when a reproduction instruction of said second speech sound data sequence is inputted during the reproduction, or after an interruption, of said first speech sound data sequence and of reproducing the read out speech sound

data;

the second step of reading out speech sound data of a section in said first speech sound data sequence corresponding to the section in said second speech sound data sequence during reproduction from said medium based on said recorded position identification data when a reproduction instruction of said first speech sound data sequence is inputted during the reproduction, or after an interruption, of said second speech sound data sequence and of reproducing the read out speech sound data; and

the third step of reading out speech sound data of a section group in said third speech sound data sequence corresponding to the section in said first or second speech sound data sequence during reproduction from said medium based on said recorded position identification data when a reproduction instruction of said third speech sound data sequence is inputted during the reproduction, or after an interruption, of said first or second speech sound data sequence and of reproducing the read out speech sound data.

23. The data reproduction system according to Claim 22, wherein said first step is the step of shifting back the position of the speech sound data in said second speech sound data sequence that has been read out from said medium by an

instructed amount in units of a notch with respect to said sections when a return instruction is inputted at the time of input of said reproduction instruction.

24. The data reproduction system according to Claim 22, wherein said second step is the step of shifting back the position of the speech sound data in said first speech sound data sequence that has been read out from said medium by an instructed amount in units of a notch with respect to said sections when a return instruction is inputted at the time of input of said reproduction instruction.

25. The data reproduction system according to Claim 22, wherein said third step is the step of shifting back the position of the speech sound data in said third speech sound data sequence that has been read out from said medium by an instructed amount in units of a notch with respect to said section groups when a return instruction is inputted at the time of input of said reproduction instruction.

26. A data reproduction system for reproducing speech sound data and text data that have been recorded on a medium in advance, wherein

said speech sound data and text data include at least: a first speech sound data sequence divided into a plurality of sections; a second speech sound data sequence which has

contents corresponding to said first speech sound data sequence, which is made up of different speech sound data and which is divided into a plurality of sections corresponding to the sections of said first speech sound data sequence; a text data sequence divided into a plurality of sections corresponding to the sections of said first speech sound data sequence which is made up of text data having contents corresponding to said first or second speech sound data sequence; and recorded position identification data showing the recorded positions of said first and second speech sound data sequences as well as a text data sequence, respectively, in said medium using each of said sections, and

the data reproduction system comprises at least:

the first step of reading out speech sound data of a section in said second speech sound data sequence corresponding to the section in said first speech sound data sequence during reproduction from said medium based on said recorded position identification data when a reproduction instruction of said second speech sound data sequence is inputted during the reproduction, or after an interruption, of said first speech sound data sequence and of reproducing the read out speech sound data;

the second step of reading out speech sound data of a section

in said first speech sound data sequence corresponding to the section in said second speech sound data sequence during reproduction from said medium based on said recorded position identification data when a reproduction instruction of said first speech sound data sequence is inputted during the reproduction, or after an interruption, of said second speech sound data sequence and of reproducing the read out speech sound data; and

the third step of reading out text data of the corresponding section in said text data sequence from said medium based on said recorded position identification data during the reproduction of, or after an interruption of, said first or second speech sound data sequence and of displaying the read out text data.

27. The data reproduction system according to Claim 26, wherein said first step is the step of shifting back the position of the speech sound data in said second speech sound data sequence that has been read out from said medium by an instructed amount in units of a notch with respect to said sections when a return instruction is inputted at the time of input of said reproduction instruction.

28. The data reproduction system according to Claim 26, wherein said second step is the step of shifting back the

position of the speech sound data in said first speech sound data sequence that has been read out from said medium by an instructed amount in units of a notch with respect to said sections when a return instruction is inputted at the time of input of said reproduction instruction.

3. DETAILED DESCRIPTION OF THE INVENTION

[Field of the Invention]

The present invention relates to a data recording medium such as a CD-ROM, a DAT, or the like, in which speech sound data, and the like, is recorded, and relates to a data reproduction system for reproducing speech sound data, and the like, that has been recorded on such a medium in advance.

[Prior Arts]

A variety of speech sound data recorded on a medium such as a cassette tape have been provided for the purpose of self-study of a language such as English conversation, the practice of reciting Chinese poems, the study of law, etc. Here, a cassette tape for the self-study of English conversation will be described as an example. A sequence of English vocal sound (speech sound data), for example, is recorded on a cassette tape (or a record); and such a tape is combined with text books as teaching materials and is used by the learner for self-

study according to the prior art. Thus there are a variety of levels starting from novice to expert.

It is possible to listen to the same portion of the tape again by slightly rewinding it, in a case where the portion cannot be heard while listening to the tape according to the prior art. In addition, it is also possible to listen to the tape many times repeatedly. In the study of a foreign language, however, there are some portions that cannot be heard even in a case where those portions are listened to many times. In such a case there is no other method than referring to the text book according to the prior art. This is because in many cases some portions cannot be heard even if the tape is played slowly when a language, for example American English, is vocalized in a manner in which a plurality of words is pronounced as one word.

However, primarily, the following two problems exist concerning the manner of understanding the portions that cannot be heard on a tape as described above by referring to a text book.

The first problem is the habit of modern Japanese people who rely on text due to an extremely low listening ability, in comparison with their reading ability, leading to a tendency where listening ability does not improve over time. As such, it is desirable not to read text as much as possible during

listening practice. According to a practice system using a conventional style tape (or other media), however, there is no other appropriate method than to read a text book in the case of encountering a situation as described above. There are only a few kinds of these negative effects in the case where it becomes necessary to read a textbook approximately once every 5 or 10 minutes; on the other hand, such materials may be too easy for that person so as not to provide good practice time. In the case where teaching materials appropriate for practicing are selected, many portions that cannot be heard are encountered. Reading the textbook at such times prevents the person from eliminating the bad habit of relying on reading text.

The second problem is the trouble of reading a textbook whenever it is necessary to do so. In particular, it is desirable to study without reading a textbook when practicing one's listening skills on a train or the like.

Therefore, the two systems described below can be considered in the case wherein the problem described above of referring to a textbook is attempted to be solved according to the prior art.

The first system is to sequentially record the English speech sound data at a natural speed, spoken by a native speaker and

English speech sound data of the same content spoken slowly by separating the data word by word so that it is easy for a Japanese person to listen to; and to sequentially record additional Japanese speech sound data for explaining this English speech sound data in Japanese if necessary. This is described concretely as follows: first "It's not much of a problem. I'd second that." is recorded at a length of 1.8 seconds at a native speaker speed and then "It is not much of a problem. I would second that." is recorded at a length of 5.0 seconds at a slower speed; and next the following is recorded at a length of 37 seconds in Japanese: "'It' indicates a preceding 'something' and 'not much of a problem' is an idiom meaning 'no problem' or 'no worries.' 'I'd' is the abbreviated version of 'I would' while 'I could' may be abbreviated to 'I'd' in the same manner. The meaning of the word 'second' is to 'indicate' or to 'agree,' so that 'I'd second that' means 'I agree with that.'"

The above is sequentially recorded and therefore is sequentially reproduced unless the operator skips or fast forwards to different parts of the media.

The second system is to record the English speech sound data as described above spoken by a native speaker at a natural speed on, for example, the first track of the tape; then to record

the English speech sound data spoken slowly by separating the data word by word on the second track; and finally to record additional Japanese explanation on another track if necessary. According to this system speech sound data of a desired content can be reproduced by designating the track to be reproduced. [Problems to be Solved by the Invention]

There are novices and experts from among people who use English conversation tapes. Therefore, the following problem arises when using the first system described above. That is, to say, for the experts the primary purpose is to listen to English spoken by a native speaker and therefore it becomes troublesome to listen to English spoken slowly and the Japanese explanation because it interrupts the English spoken by a native speaker. Thus, though one tries to select and reproduce only the parts of English spoken by a native speaker it is not easy to find them due to the sequential recording leading to the necessity of complicated switching operations.

On the other hand, English spoken by a native speaker is too fast to be listened to for novices and it becomes necessary to listen to slowly spoken English and Japanese explanation to be compared to the English spoken by a native speaker. However, the details cannot be fully understood when using the first system where the sentences are heard sequentially, are

long or include a plurality of parts that cannot be heard. Therefore, when the parts spoken by a native speaker are attempted to be heard in comparison to the parts spoken slowly by separating the data word by word and the parts of the Japanese explanation, it becomes difficult to find the different parts due to the sequential recording, leading to the necessity of very complicated switching operations.

In comparison to this, when using the second system described above the respective parts of the speech sound data are recorded separately on different tracks as described above and therefore it is comparatively easy to sequentially reproduce, for example, only the parts spoken by a native speaker. However, it is difficult to find the recorded position on the second track that corresponds to the part that one could not hear in the case where, for example, the part of English spoken by a native speaker on the first track could not be heard. Of course, a correspondence table between the tape counter number of the first track and the tape counter number of the second track can be prepared so that in the case where English by a native speaker cannot be heard in a certain recorded position of the first track, the tape can be fast-forwarded or rewound to the corresponding recorded position of the second track according to the above described correspondence table and one can listen

to slowly spoken English. In addition, it is not impossible to automate this process. However, there still are some drawbacks such as the length of time needed for fast-forwarding or rewinding the tape.

Therefore, a conventional data recording medium or data reproduction system for self-study learners of English conversation cannot be provided in a manner where experts and novices are both satisfied. Similar problems arise in materials for practicing recitation of Chinese poems and the study of law or other such subjects in addition to the self-study of a language.

Thus, an object of the present invention is to provide a data recording medium and a data reproduction system that provide appropriate speech sound data, etc., for both experts, that is to say, those who have mastered a field of study, and novices, that is to say, those who have just begun study, so that they both can be satisfied.

[Means for Solving Problems]

The first mode of a data recording medium according to the present invention is a medium (for example a CD-ROM, or DAT) in which speech sound data is recorded and is characterized by being provided with at least the following three regions. That is, to say that, at the least, three regions are provided

as follows:

the first region for recording a first speech sound data sequence (for example, English speech sound data sequence spoken by a native speaker) divided into a plurality of sections;

the second region for recording a second speech sound data sequence (for example, English speech sound data sequence spoken by separating the data word by word) which has contents corresponding to the first speech sound data sequence, is made up of different speech sound data and is divided into a plurality of sections corresponding to the plurality of sections of the first speech sound data sequence; and

the third region for recording, as a directory for example, recorded position identification data which indicates the respective recorded positions of the first and second speech sound data sequences in the medium using each of the sections.

The second mode of the data recording medium according to the present invention is a medium (for example a CD-ROM or DAT) in which speech sound data is recorded and is characterized by being provided with at least the following four regions. That is, to say that, at the least, four regions are provided as follows:

the first step of recording a first speech sound data

sequence (for example English speech sound data sequence spoken by a native speaker) divided into a plurality of sections;

the second region for recording a second speech sound data sequence (for example, English speech sound data sequence spoken by separating the data word by word) which has contents corresponding to the first speech sound data sequence, is made up of different speech sound data and is divided into a plurality of sections corresponding to the sections of the first speech sound data sequence;

the third region for recording a third speech sound data sequence (for example, a speech sound data sequence explained in Japanese) which has contents corresponding to the first and second speech sound data sequences, is made up of different speech sound data and is divided into a plurality of section groups, each group including one or more sections of the first and second speech sound data sequences put together collectively; and

the fourth region for recording, as a directory for example, recorded position identification data which indicates the recorded positions of the first and second speech sound data sequences, respectively, using each of the sections and for recording, as a directory for example, recorded position identification data which indicates the recorded positions of

the third speech sound data sequence using each of the section groups.

The third mode of the data recording medium according to the present invention is a medium (for example a CD-ROM or DAT) in which speech sound data is recorded and is characterized by being provided with at least the following four regions. That is, to say that, at the least, four regions are provided as follows:

the first region for recording a first speech sound data sequence (for example English speech sound data sequence spoken by a native speaker) divided into a plurality of sections;

the second region for recording a second speech sound data sequence (for example, English speech sound data sequence spoken by separating the data word by word) which has contents corresponding to the first speech sound data sequence, is made up of different speech sound data and is divided into a plurality of sections corresponding to the sections of the first speech sound data sequence;

the third region for recording a text data sequence which is made up of text data having contents corresponding to the first or second speech sound data sequence which is divided into a plurality of sections corresponding to the sections of the first speech sound data sequence; and

the fourth region for recording, as a directory for example, recorded position identification data which indicates the recorded positions of the first and second speech sound data sequences as well as the text data sequence, respectively, using each of the sections.

In addition, the first mode of the data reproduction system according to the present invention is a data reproduction system for reproducing speech sound data that has been recorded on a medium (for example, a CD-ROM, or a DAT) in advance, and is characterized in that the speech sound data includes at least the following three types of data, and in that the data reproduction system is provided with the below described two steps. That is to say,

the speech sound data at least includes: a first speech sound data sequence (for example, English speech sound data sequence spoken by a native speaker) divided into a plurality of sections; a second speech sound data sequence (for example, English speech sound data sequence spoken by separating the data word by word) which has contents corresponding to the first speech sound data sequence, is made up of different speech sound data and is divided into a plurality of sections corresponding to the sections of the first speech sound data sequences; and recorded position identification data (for example, data

recorded as a directory) which indicates recorded positions of the first and second speech sound data sequences, respectively, in the medium using each of said sections, and

the data reproduction system is provided with at least the two steps as follows:

the first step of reading out speech sound data of a section in the second speech sound data sequence corresponding to the section in the first speech sound data sequence during reproduction or after an interruption from the medium based on the recorded position identification data when a reproduction instruction of the second speech sound data sequence is inputted during the reproduction, or after an interruption, of the first speech sound data sequence, and of reproducing the read out speech sound data; and

the second step of reading out speech sound data of a section in the first speech sound data sequence corresponding to the section in the second speech sound data sequence during reproduction from the medium based on the recorded position identification data when a reproduction instruction of the first speech sound data sequence is inputted during the reproduction, or after an interruption, of the second speech sound data sequence, and of reproducing the read out speech sound data.

The second mode of the data reproduction system according to the present invention is a data reproduction system for reproducing speech sound data that has been recorded on a medium (for example, a CD-ROM, or a DAT) in advance, and is characterized in that the speech sound data includes at least the following four types of data, and in that the data reproduction system is provided with the below described three steps. That is to say,

the speech sound data at least includes: a first speech sound data sequence (for example, English speech sound data sequence spoken by a native speaker) divided into a plurality of sections; a second speech sound data sequence (for example, English speech sound data sequence spoken by separating the data word by word) which has contents corresponding to the first speech sound data sequence, is made up of different speech sound data and is divided into a plurality of sections corresponding to the sections of the first speech sound data sequence; a third speech sound data sequence (for example, a speech sound data sequence explained in Japanese) which has contents corresponding to the first and second speech sound data sequences, is made up of different speech sound data and is divided into a plurality of section groups, each group made up of one, or more, sections from among the above described

sections; and recorded position identification data (for example, data recorded as a directory) which indicates the recorded positions of the first and second speech sound data sequences, respectively, in the medium using each of the sections and which indicates the recorded position of the third speech sound data sequence in the medium using each of the section groups, and

the data reproduction system is provide with at least the three steps as follows:

the first step of reading out speech sound data of a section in the second speech sound data sequence corresponding to the section in the first speech sound data sequence during reproduction from the medium based on the recorded position identification data when a reproduction instruction of the second speech sound data sequence is inputted during the reproduction, or after an interruption, of the first speech sound data sequence, and of reproducing the read out speech sound data;

the second step of reading out speech sound data of a section in the first speech sound data sequence corresponding to the section in the second speech sound data sequence during reproduction from the medium based on the recorded position identification data when a reproduction instruction of the

first speech sound data sequence is inputted during the reproduction, or after an interruption, of the second speech sound data sequence, and of reproducing the read out speech sound data; and

the third step of reading out speech sound data of a section group in the third speech sound data sequence corresponding to the section in the first or second speech sound data sequence during reproduction from the medium based on the recorded position identification data when a reproduction instruction of the third speech sound data sequence is inputted during the reproduction, or after an interruption, of the first or second speech sound data sequence and of reproducing the read out speech sound data.

The third mode of the data reproduction system according to the present invention is a data reproduction system for reproducing speech sound data and text data that have been recorded on a medium (for example, a CD-ROM, or a DAT) in advance, and is characterized in that the speech sound data and text data include at least the following four types of data, and in that the data reproduction system is provided with the below described three steps. That is to say,

the speech sound data and text data include at least: a first speech sound data sequence (for example, English speech sound

data sequence spoken by a native speaker) divided into a plurality of sections; a second speech sound data sequence (for example, English speech sound data sequence spoken by separating the data word by word) which has contents corresponding to the first speech sound data sequence, is made up of different speech sound data and is divided into a plurality of sections corresponding to the sections of the first speech sound data sequence; a text data sequence divided into a plurality of sections corresponding to the sections of the first speech sound data sequence which is made up of text data having contents corresponding to the first or second speech sound data sequence; and recorded position identification data (for example, data recorded as a directory) indicating the recorded positions of the first and second speech sound data sequences as well as the text data sequence, respectively, in the medium using each of the sections, and the data reproduction system is provided with at least three steps as follows:

the first step of reading out speech sound data of a section in the second speech sound data sequence corresponding to the section in the first speech sound data sequence during reproduction from the medium based on the recorded position identification data when a reproduction instruction of the

second speech sound data sequence is inputted during the reproduction, or after an interruption, of the first speech sound data sequence, and of reproducing the read out speech sound data;

the second step of reading out speech sound data of a section in the first speech sound data sequence corresponding to the section in the second speech sound data sequence during reproduction from the medium based on the recorded position identification data when a reproduction instruction of the first speech sound data sequence is inputted during the reproduction, or after an interruption, of the second speech sound data sequence, and of reproducing the read out speech sound data; and

the third step of reading out text data of the corresponding section in the text data sequence from the medium based on the recorded position identification data during the reproduction of, or after an interruption of, the first or second speech sound data sequence, and of displaying the read out text data on an LCD, or the like.

[Action]

The first mode of the data recording medium according to the present invention is configured as described above and therefore the first region for recording the first speech sound

data sequence serves to record the basic speech sound data by dividing the first speech sound data into a plurality of sections based on, for example, breaks concerning pronunciation, breaks concerning some linguistic elements or without having particular relationships with these breaks; the second region for recording the second speech sound data sequence serves to record speech sound data that has been, for example, paraphrased from the speech sound data of the first speech sound data sequence by dividing the second speech sound data into a plurality of sections; the third step of recording the recorded position identification data serves to identify the recorded positions of the above described two data sequences in the medium using each of the sections.

The second mode of the data recording medium according to the present invention is configured as described above and therefore the first region for recording the first speech sound data sequence serves to record the basic speech sound data by dividing the first speech sound data into a plurality of sections based on, for example, breaks concerning pronunciation, breaks concerning some linguistic elements or without having particular relationships with these breaks; the second region for recording the second speech sound data sequence serves to record speech sound data that has been, for

example, paraphrased from the speech sound data of the first speech sound data sequence by dividing the second speech sound data into a plurality of sections; the third region for recording the third speech sound data sequence serves to record speech sound data for individual explanations or a unified explanation, or the like, of a plurality of sections of the first and second speech sound data sequences; and the fourth region for a recording the recorded position identification data serves to identify recorded positions of the above described three data sequences in the medium using each of the sections or each of the section groups.

The third mode of the data recording medium according to the present invention is configured as described above and therefore the first region for recording the first speech sound data sequence serves to record the basic speech sound data by dividing the first speech sound data into a plurality of sections based on, for example, breaks concerning pronunciation, breaks concerning some linguistic elements or without having particular relationships with these breaks; the second region for recording the second speech sound data sequence serves to record speech sound data that has been, for example, paraphrased from the speech sound data of the first speech sound data sequence by dividing the second speech sound

data into a plurality of sections; the third region for recording a text data sequence serves to record text data corresponding to the first or second speech sound data sequence by dividing the text data into a plurality of sections; and the fourth region for recording the recorded position identification data serves to identify the recorded positions of the above described three data sequences in the medium using each of the sections.

In addition, the first mode of the data reproduction system according to the present invention is configured as described above and, therefore, the first and second steps serve to switch the reproduction of a speech sound data sequence into that of the corresponding section of another speech sound data sequence that has been instructed, wherein the recorded position identification data serves to indicate the recorded position of the switched data sequence using each of the sections at the time of the above described switch.

The second mode of the data reproduction system according to the present invention is configured as described above and, therefore, the first, second, and third steps serve to switch the reproduction of a speech sound data sequence into that of the corresponding section or section group of another speech sound data sequence that has been instructed, wherein the

recorded position identification data serves to indicate the recorded position of the switched data sequence using each of the sections, or each of the section groups, at the time of the above described switch.

The third mode of the data reproduction system according to the present invention is configured as described above and, therefore, the first and second steps serve to switch the reproduction of a speech sound data sequence into that of the corresponding section of another speech sound data sequence that has been instructed and the fourth step serves to display the text data in the section corresponding to the speech sound data during reproduction, wherein the recorded position identification data serves to indicate the recorded position of the data sequence that is displayed or to which the reproduction is switched at the time of this displaying and of this switching of the reproduction.

[Preferred Embodiment]

In the following, the basic contents of the above described six modes of the present invention are respectively described prior to concrete descriptions of the embodiments of the present invention. Here, the respective constituent features described in the claims are, of course, not limited by the basic contents.

The constituent features of the first mode of the data recording medium according to the present invention are clarified in Claim 1. A recording medium of this mode is characterized by recording at least two speech sound data sequences. That is to say, the first speech sound data sequence is, for example, made up of English speech sound data spoken by a native speaker at a natural speed and this speech sound data sequence is divided into a plurality of sections (so called segments). Though the second speech sound data sequence corresponds to the contents of the above described first speech sound data sequence, it is made up of different speech sound data and is, for example, English speech sound data spoken at a slower speed by separating the data word by word.

The important thing here is that the above described first and second speech sound data sequences are divided respectively into pluralities of sections so that the respective sections correspond to each other. In the case wherein the t'th section of the first speech sound data sequence is "It's" spoken by a native speaker, for example, the t'th section of the second speech sound data sequence becomes "It is" spoken by separating the data word by word. Accordingly, the sections in the first and second speech sound data sequences mean segments divided based on breaks concerning pronunciation or concerning some

linguistic elements. Here, the formation of sections based on such breaks is not essential to the present invention, but rather, the sections may be based on time intervals irrelevant of the above described breaks as shown in the below example using a DAT. In addition, the expression "having contents corresponding to the second speech sound data sequence and being made up of different speech sound data" can be stated in other words where the second speech sound data sequence has the same meaning as the first speech sound data sequence as long as the language is related and has a different pronunciation from the first speech sound data sequence.

Furthermore, the recording medium of this mode is characterized by recording recorded position identification data. Thus, this recorded position identification data functions so as to indicate at which position in the medium the speech sound data of the first and second speech sound data sequences is recorded. Accordingly, the position in the medium in which "It is" in the second speech sound data sequence corresponding to "It's" of the t'th section in the first speech sound data sequence, for example, is recorded can be recognized according to the above described recorded position identification data.

As a result, it can be understood that the first and second

speech sound data sequences and the recorded position identification data are not recorded irrelevant of each other but rather are recorded to maintain a certain relationship so that the respective pieces of data are combined in an organic manner using the sections as units. That is to say, the first and second speech sound data sequences become a pair with each other wherein the recorded position identification data makes these sequences correspond to each other by the respective sections. Here, the recording medium of this mode is the most basic mode of the present invention.

The constituent features of the second mode of the data recording medium according to the present invention are clarified in Claim 7. Thus, this mode differs from the above described first mode in the point wherein the third speech sound data sequence is recorded on the medium in addition to the first and second speech sound data sequences according to this mode.

The important thing here is that the third speech sound data sequence is divided into section groups wherein each section group is made up of one, or more, of the sections of the first and second speech sound data sequences. In other words, one section group of the third speech sound data sequence includes one, or more, of the sections of the first and second speech sound data sequences and accordingly one section group pairs

up with one or more sections.

Furthermore, a recording medium of this mode has an effect wherein the recorded position identification data indicates the recorded positions of the contents of the third speech sound data sequence by using each of the section groups. Accordingly, the first, second and third speech sound data sequences as well as the recorded position identification data are recorded on the medium so as to maintain a certain relationship with each other wherein the respective pieces of the data are combined in an organic manner using the sections and the section groups as units.

The constituent features of the third mode of the data recording medium according to the present invention are clarified in Claim 13. Thus, this mode differs from the above described second mode in the point wherein a text data sequence is recorded on the medium in addition to the first and second speech sound data sequences according to this mode. This text data sequence is made up of text data having contents corresponding to the first and second speech sound data sequences and represents English spoken by a native speaker with text, for example.

This text data sequence is also divided into sections corresponding to the section of the first and second speech

sound data sequences. In addition, the recorded position identification data functions such that the recorded position of this text data sequence is identified using each of the sections. Accordingly, the first and second speech sound data sequences and the text data sequence correspond to each other using the sections as units. Here, in the case wherein the third speech sound data sequence according to the second mode is added as recorded data to the recording system of this mode, one or more sections of the first and second speech sound data sequences as well as the text data sequence correspond to one section group of the third speech sound data sequence.

The modes of the data recording medium according to the present invention are described above. In comparison to this, the data reproduction system according to the present invention has the following three modes and these modes are compared to the three modes of the data recording medium respectively shown above.

The constituent features of the first mode of the data reproduction system according to the present invention are shown in Claim 19. The reproductive system of this mode presupposes that speech sound data according to the first mode of the recording medium has been recorded on the medium in advance. Furthermore, it is characterized by being provided

with at least the following two steps.

The first step is switching from the reproduction of the first speech sound data sequence to the second speech sound data sequence. Thus, this switching is carried out using the sections as units. In the case wherein a reproduction instruction of the second speech sound data sequence is inputted when the t'th section of the first speech sound data sequence is being reproduced, for example, the t'th section of the second speech sound data sequence is read out based on the recorded position identification data so as to reproduce the speech sound data thereof. The second step is switching from the reproduction of the second speech sound data sequence to the first speech sound data sequence and this switching is also carried out according to the sections in the same manner as the first step.

Here, the first and second steps are not limited to the above description, but rather, a variety of modifications thereof are possible. The so called return instruction is a representative example. That is to say, in the case wherein a return instruction is inputted after the reproduction is temporarily interrupted due to a stop instruction during the reproduction, the readout position of the speech sound data is shifted back by the instructed amount and thereby

reproduction of the speech sound data is carried out in the desired manner.

The constituent features of the second mode of the data reproduction system according to the present invention are shown in Claim 22. Thus, this mode differs from the above described first mode in the point that this mode presupposes that the speech sound data according to the second mode of the recording medium has been recorded on the medium in advance and that this mode is provided with the third step of switching from the reproduction of the first or second speech sound data sequence to the third speech sound data sequence.

In the case when "It's" by a native speaker cannot be heard during the reproduction of the first speech sound data sequence, for example, the reproduction is switched to the second speech sound data sequence according to the first step so that "It is" spoken slowly by separating the data word by word can be heard. Thus, in the case where the meaning and the grammar of this are desired to be known in Japanese, the reproduction may be switched to the third speech sound data sequence according to the third step.

The system of this mode can, of course, be modified so that a return instruction and a stop instruction as described in the above first mode can be combined and used.

The constituent features of the third mode of the data reproduction system according to the present invention are shown in Claim 26. Thus, this mode differs from the above described second mode in the point that this mode presupposes that the speech sound data and the text data according to the third mode of the recording medium have been recorded on the medium in advance and the text data sequence is displayed during the reproduction of the first or second speech sound data sequence according to this mode.

In the case where "It's" of the first speech sound data sequence is reproduced, for example, "It's" or "It is" is displayed in a predetermined display part. Here, it is not necessary for this display to completely synchronize with the speech sound data during reproduction chronologically and text may be displayed with a slight delay or may be displayed slightly ahead of time.

Next, several embodiments of the present invention are concretely described in reference to the attached drawings.

First, the first example of a recording medium according to the present invention is described in reference to Fig. 1 through Fig. 3. Fig. 1 is a view showing data sequences A, B and C as well as the recorded contents thereof where a recording medium of the present invention is applied for self-study of

English conversation. In this figure, data sequence A indicates an English speech sound data sequence (first speech sound data sequence) spoken by a native speaker and is divided into a plurality of segments 621 to 627 indicating the sections thereof. Data sequence B indicates an English speech sound data sequence (second speech sound data sequence) spoken slowly by separating the data word by word and this is speech sound data of sections made up of English words and phrases corresponding to segments 621 to 627. Data sequence C indicates a speech sound data sequence (third speech sound data sequence) for Japanese explanation and corresponds to segments 621 to 624 and 625 to 627 of data sequences A and B, respectively.

Fig. 2 is a view showing the relationship between time and capacity of each segment in the example shown in Fig. 1. In this figure, one second corresponds to a capacity of 6 kilobytes. It takes 0.4 seconds for "much of a" of segment 623 to be spoken at a native speaking speed, for example, and therefore, the recording capacity used in the medium is 2.4 kilobytes.

Fig. 3 is a view showing one example of a directory in the example shown in Fig. 1 and Fig. 2. In this figure, one segment of the directory is formed of $9 \times 3 = 27$ bytes. Data sequences A, B and C correspond data sequences A, B and C of Fig. 1 respectively. In addition, C of one byte indicates an attribute

and when this is 129, that is to say "10000001" according to the bit expression (8 bits), this indicates that it has the same contents as the contents of the previous data sequence.

Position data M, S and B (one byte each) is a parameter showing the positions in a CD-ROM that has been standardized in the industry. That is to say M indicates minutes, S indicates seconds, and B indicates blocks, respectively, wherein one block is made up of 2048 bytes. Thus, 75 blocks form an amount for one second. Accordingly, the maximum numbers are M = 59, S = 59, and B = 74. SB of the next two bytes indicates the start byte and LLL of the following three bytes indicates the byte length. Here, minutes and seconds are used as parameters indicating position, because the CD-ROM was originally developed for music recording, and therefore, the recorded position is represented as a length of a time from the start. Thus, in the case of a CD-ROM, minutes and seconds are not at all related to the period of time of reproduction, but rather this data simply represents the recorded position.

As a result, as for "much of a" of segment 623, for example, English spoken by a native speaker is recorded starting from the 1354th byte of the 0 minutes 11 seconds 42 blocks with a length of 2400 bytes, English spoken slowly by separating the data word by word is recorded starting from the 1706th byte of

the 0 minutes 11 seconds 43 blocks with a length of 7200 bytes and Japanese explanation is recorded starting from the 1282nd byte of 0 minutes 11 seconds 6 blocks with a length of 72000 bytes. Here, segment numbers 621 to 627 are not in the memory but correspond to the addresses thereof.

Next, effects of the recording medium shown in the above described Fig. 1 through Fig. 3 are described.

As shown in Fig. 3, the region from the 826th byte to the $(826 + 1200 - 1 =) 2025^{\text{th}}$ byte at 0 minutes 11 seconds 3 blocks in the medium serves to record the data sequence wherein the segment is 621 and attribute C is 0, that is to say, "It's" spoken by a native speaker. In addition, the region starting from the 2026th byte to the $(2026 + 5400 - 1 =) 7425^{\text{th}}$ byte at 0 minutes 11 seconds 3 blocks in the medium, that is to say, to the 1282nd byte at 0 minutes 11 seconds 6 blocks serves to record the data sequence wherein the segment is 621 and attribute C is 64, that is to say, "It is" spoken slowly by separating the data word by word. In addition, the region from the 1282nd byte at 0 minutes 11 seconds 6 blocks to 1601st byte at 0 minutes 11 seconds 41 blocks in the medium serves to record the data sequence wherein the segment is 621 and attribute C is 128, that is to say, the Japanese explanation.

As described above, in the case where a directory as shown

in Fig. 3 is formed, the data sequence shown in Fig. 1 can be recorded to have the capacity and reproduction period of time as shown in Fig. 2.

Next, a reproduction system for the recording medium described in reference to Fig. 1 through Fig. 3 is described in reference to Fig. 4 through Fig. 9.

Fig. 4 is a block diagram showing the configuration of one example of a reproduction unit. In this figure, a medium 1 is, for example, a CD-ROM and is set in a reproduction mechanism 2. Reproduction mechanism 2 is connected to a CPU 5 via a disk interface (I/F) 3 and a bus 4. In addition, a ROM 6 of 32 kilobytes, for example, for storing a program and a RAM 7 of 256 kilobytes, for example, for temporarily storing a directory or a speech sound data sequence, are connected to bus 4. Furthermore, bus 4 is connected to a handset interface (I/F) 9 for sending and receiving data to and from a handset 8 for manual operation and is connected to a D/A converter 12 that is connected to an external terminal 11 and handset 8 via an AMP 10. Here, an earphone 13 is connected to handset 8.

Fig. 5(a) and Fig. 5(b) are diagrams showing memory allocations of ROM 6 and RAM 7. As shown in Fig. 5(a), a program of 32 kilobytes is stored in ROM 6. As shown in Fig. 5(b), RAM 7 is allocated for a buffer of 50 blocks for $(50 + 50 =) 100$

kilobytes, a directory for $(75 + 75 =)$ 150 kilobytes and a system area for 6 kilobytes. Accordingly, a speech sound data sequence of 50 blocks is always stored in RAM 7 and a directory (only the portion of data sequence A takes up approximately 30 minutes) for 150 kilobytes + 27 bytes = 5555 segments is also stored in RAM 7.

Here, a CD-ROM is used as the medium in the above described example and the capacity of a representative CD-ROM is 552 megabytes. Units of minutes, seconds, and blocks are used to represent the address in the CD-ROM. As described above, one block is formed of 2048 bytes, 75 blocks form one second and 60 seconds form one minute while the CD-ROM holds the maximum of 59 minutes 59 seconds 74 blocks. Accordingly, the maximum storage amount becomes $(2048 \times 75 \times 60 \times 60 =)$ 552.96 megabytes. The first two seconds out of the above maximum storage amount are used for formatting the CD-ROM and cannot be used by the user and thereby, the maximum capacity becomes 552.6528 MB to be precise. In the case wherein a directory is allocated to a portion corresponding to the first 20 seconds out of the maximum capacity, a directory capacity of 3 megabytes can be secured in the CD-ROM.

Next, a calculation example concerning the capacity is shown.

A sound sample is produced consisting of 16 kilo-samples/second according to an ADPCM system and 3 bits are used for one sample. This leads to 48 kilobits/second and accordingly, 6 kilobytes/second, and therefore, it is necessary for the transfer rate to be adjusted to this rate. Here, in the case of 16 kilo-samples/second, f characteristics up to 8 KHz are exhibited. Accordingly, consonants can be recorded sufficiently.

According to the above described presumption, a capacity of $(6 \text{ kilobytes} \times 3600 \text{ seconds} =) 21.6 \text{ megabytes}$ is required for recording sound for one hour. In general, 552 megabytes are recorded on one CD-ROM, including error corrections. The CD-ROM, excluding the directory portion, can store 549 megabytes of speech sound data. Accordingly, it becomes possible to record speech sound data for $(549 + 21.6 =) 25 \text{ hours } 24 \text{ minutes}$. Thus, in the case where the CD-ROM is used for the study of English conversation, supposing that a one hour story spoken by a native speaker at a natural speed is provided, the portion of the story that is spoken slowly by separating the data word by word would take four hours, four times longer than the natural speed. Thus, the total hours become 20 hours even in the case where 15 total hours are required for the explanation portion.

Next, the method of dividing the one hour conversation portion into several sections (segments) is considered and then presuming that each second is made up of four segments on average, the one hour would be divided into 14400 segments. 27 bytes are required for a directory for one segment, and therefore, approximately 389 kilobytes are required in total which can be sufficiently stored in the above described storage place for directories of 3 megabytes, and therefore, all of the directories for the one hour story can be stored.

Fig. 6 is a detailed illustration of handset 8 and earphone 13 shown in Fig. 4. The front side of handset 8 is provided with a segment display part 21 for displaying the segment number, an FF switch 22 for fast-forwarding, A, B and C instruction switches 23A, 23B and 23C for instructing the respective reproductions of speech sound data sequences A, B and C, a STOP button 24 for instructing the stoppage of the reproduction, and a REV button 25 for instructing the return of reproduction. In addition, earphone 13 is connected to handset 8 via a cord 26 and handset 8 is connected to the body via a cord 27.

Fig. 7(a) is a perspective view of the above described reproduction unit and Fig. 7(b) is a perspective view of this reproduction unit being contained in a case 29 to which a strap 28 is attached. The CD-ROM has been miniaturized and reduced

in weight as a result of the recent progress in technology, and therefore, can be made portable as shown in the figures. Fig. 8 shows the condition where the handset is hand held.

Next, the effects of the present embodiment are described in reference to one example of the reproduction sequence according to the reproduction system above described. Fig. 9 is the sequence diagram thereof. Fig. 9(a) shows the flow of data when the portion of data sequence A, that is to say, the native speaker English, is continuously heard. At this time button 23A in Fig. 6 is pressed so that the reproduced sound is heard as is.

Next, the sequence of Fig. 9(b) is described. First, data sequence A of segment 621 is reproduced. Here, when "much of a" of segment 623 cannot be heard well, STOP button 24 is immediately pressed. At this time, the segment number of segment display part 21 has become 624. Thus, REV button 25 is pressed only once. Here, when REV button 25 is held down the segment continuously moves backwards and every time REV button 25 is pressed the segment of display part 21 moves backwards by one. When REV button 25 is pressed once under the condition where the segment is 624, the segment of display part 21 becomes 623.

Next, when switch 23B is pressed "much/ of/ a" is heard spoken

slowly by separating the data word by word. When the unit is left as is, it proceeds to "problem" through data sequence B. When "much/ of/ a" is heard, STOP button 24 is pressed and REV button 25 is again pressed so that the segment becomes 623. Then, when switch 23C is pressed the Japanese explanation data sequence C can be heard.

Some portions of the explanation relate to phrases made up of several words and therefore several segment numbers may be covered by one explanation. When the lowest level bit of the byte showing attribute C is 1, this segment (where attribute C is 129) shows the attribute with a meaning of the same content as the previous segment number of the same data sequence. When such a segment is encountered, it can be skipped.

Next, the second example of a recording medium and reproduction system according to the present invention is described in reference to Fig. 10 through Fig. 12. Fig. 10 is a view showing directly wherein one segment is formed of 12 bytes. That is to say, attribute C of one byte, segment number SS of two bytes, suffix number N of one byte, minute M of one byte, second S of one byte, block B of one byte, start byte SB of two bytes and byte length LLL of three bytes add up to 12 bytes.

In addition, when the first bit (top level bit) is 1, it

indicates the start of the segment and when the first bit is 0, it indicates another condition concerning the 8 bits attribute C. When the second and third bits are 0, it indicates that the segment is of data sequence A; when the second and third bits are 1, it indicates that the segment is of data sequence B; and when the second and third bits are 2, it indicates that the segment is of data sequence C. When the fourth bit is 0, it indicates there is no suffix; and when the fourth bit is 1, it indicates there is a suffix. When the fifth bit is 1, it indicates the segment has the same content as the previous segment; and when the fifth bit is 0 the segment does not have the same content.

Fig. 11 shows data sequences A, B, and C being divided into segments according to the directory of Fig. 10. Thus, this differs from the above described first embodiment in the point that the data sequences are divided into segments with suffixes in this embodiment. Fig. 12 is a view showing the relationship between the values of the second and third bits of attribute C which has 8 bits and the segments.

Next, a recording medium and reproduction system according to the third example of the present invention is described in reference to Fig. 13 and Fig. 14. Thus, this example differs from the first and second examples in the point wherein text

data is recorded and reproduced in addition to speech sound data.

Fig. 13 is a view showing the directory of this example. Thus this differs from the directory of Fig. 10 in the point that data of $D = 3$ is added to the second and third bits of attribute C in this example. When this bit is 3 it indicates that text data sequence D is recorded on the medium in the form of a code.

Fig. 14 is a frontal view of a handset used for the reproduction unit according to this example. Thus, this differs from the handset of Fig. 6 in the point wherein a text display part 41 formed of, for example, an LCD and a text display button 42 for turning on and off the display of text are provided with this handset. When text display button 42 is pressed in this kind of handset 8 it turns text display part 41 on and off so that the text data sequence D is displayed or is not displayed.

Next, the characteristics of the effects of this third embodiment are described.

The speed of the appearance of text data sequence D is controlled in accordance with the speed of data sequence A being spoken, that is to say, the length of LLL, during the reproduction of speech sound data sequence A. Thus, text begins to appear on display part 41 and completely finish appearing

during the time from the start of and to the end of the speech sound of the above described segment. That is to say, the text is displayed in complete synchronization with the speech sound.

Next, text data sequence, data D is displayed in synchronization with the length of the speech sound data sequence B during speech sound data sequence B's reproduction, and at this time the speech sound is reproduced for a period of time several times longer than the time of reproduction of speech sound data sequence A. In practice, it is considered to be more convenient for the user if the display of the text comes out sooner (the output of speech sound is slightly delayed) and then this requires further examination at the time when the unit is put into practice.

Next, a case wherein the data recording medium is a magnetic tape that cannot be accessed at random is described as the fourth embodiment of the present invention.

In this example, the basic mode of data recording and reproduction is the same as the above described examples and reproduction by skipping data on the magnetic tape is not easy. Therefore, it is desirable to provide a buffer for temporarily restoring data in order to make the reproduction unit practical. Concretely, the respective small units corresponding to data sequences A, B and C (three tracks) of the above described

virtual tracks are aligned in order so that the corresponding units of data sequences A, B and C are put together in this manner: $A_1, B_1, C_1, A_2, B_2, C_2, A_3, B_3, C_3 \dots$. Thus, when data sequence A is reproduced, only the A's such as $A_1, A_2, A_3 \dots$ are chosen for the reproduction and when data sequence B is reproduced, only the B's such as $B_1, B_2, B_2 \dots$ are chosen for the reproduction. The same can be carried out for data sequence C. At this time, B_1 and C_1 must be skipped in order to reproduce A_2 next to A_1 without interruption. This is not easy to do with a conventional tape, and therefore data must be brought into the buffer ahead of time from the magnetic tape. Thus, the skipping process and reproduction are carried out in the buffer.

In the case wherein this magnetic tape is a so-called DAT (digital audio tape) the following process occurs. First, speech sound of a native speaker recorded on track A is separated into certain units as sections. For example, the speech sound is separated into units of a constant period of time such as one second. On the other hand, track B records the contents spoken slowly (for example, at $1/3$ the speed of track A on average). Regions are secured on the tape so that if a region of 1 second is taken up on track A it corresponds to a region of four seconds on track B. In addition, the regions

are secured so that a region of 9.4 seconds can be taken up on track C for the explanation. The above described number of seconds here is an example and the present invention is not limited to this. In this example, 14.4 seconds are used as one unit. That is to say, three segments of tracks A, B, and C that correspond to each other are put together to make up 14.4 seconds. It should be noted that a portion of one second on track A is equivalent to one second that is cut out of the original sound data. A portion of four seconds on track B may have a content that is within four seconds, but may finish in 3 seconds, for example, with an extra second remaining. Here, a method for skipping this remaining portion (a one second portion) at the time of reproduction is described below.

An example of a rotary head type DAT is cited as follows. A system wherein two tracks are recorded or reproduced during one rotation of the rotary head is generally used in this example. Thus, speech sound data of 2,880 bytes can be recorded on one track. Accordingly, in the case where one segment is formed of 30 tracks the segment is made up of 86,400 bytes. In the case wherein the speech sound utilized in the above described embodiments is recorded in this segment at a sampling rate of 48 (Kbit/sec), $86,400 \text{ (bytes)} + 6,000 \text{ (bytes)} = 14.4 \text{ sec}$ can be recorded.

Next, the buffer is described.

When a RAM of 1 MB is utilized, 12 buffers can be provided as a result of this equation: $1,0480,576 + 86,400 = 12,135$. The 12 buffers are arranged to function in a ring form, so that two buffers are always selected for the data transfer from the tape to these buffers. Reproduction is carried out by taking out the data from the buffers. When the reproduction is stopped at an arbitrary place, the speech sound (already emitted from the speaker as a reproduced sound) before the point of time where the reproduction is stopped remains in the ten buffers. That is to say, speech sound for ten seconds of track A can be repetitively reproduced from the buffers without rewinding the tape.

In such a configuration, the 10 seconds of speech sound data that remains in the ten buffers can be easily transferred to any of the virtual tracks A, B or C in an arbitrary segment in the same manner as the above described embodiment using a CD-ROM.

Here, random access cannot be carried out like a CD-ROM and, therefore, the total period of time of the segments of the three tracks A, B and C that correspond to each other must always be within 14.4 seconds. At this time, the portions of virtual track A may be mechanically separated into constant time

intervals (for example, one second intervals). The portions of track B may be two seconds or five seconds, for example, and therefore variable. A period of time gained by subtracting periods of time of tracks A and B from 14.4 seconds is the period of time that is allocated for track C. Here, it is not necessary for both tracks B and C to record the portions that completely correspond to the portion in the segment of track A, but rather may form a coordinated unit of a number of segments in the vicinity.

A DAT has a sub-code region and therefore data of byte numbers indicating the respective segments (for every segment, 14.4 seconds/ 30 tracks) and interfaces between tracks A, B and C in the respective segments can be recorded in this sub-code region.

Next, an amount of speech sound data that can be contained in a DAT is concretely calculated.

In general, a DAT of one hour can record 240,000 tracks. One segment is made up of 30 tracks and the following 6 tracks are used as pause tracks so that the portions recording speech sound data do not become scratched even in the case wherein the rotary head passes through a great number of times when the tape is stopped. In such a configuration one unit is formed of 36 tracks so that 6,666 units can be recorded in total which

becomes $6,666 \times 14.4$ seconds = 26 hours, 39 minutes and 50 seconds. One hour out of that is allocated for track A and therefore 25 hours, 39 minutes and 50 seconds can be used for tracks B and C. Even if four hours are used for track B spoken slowly, more than 21 hours can be allocated for track C for the explanation and, therefore, the tape has a sufficient length.

Here, as for how the segments of tracks A, B and C are created, track A may be divided at places where there are some pauses according to speech sound, so that the length of the segments become approximately one second, instead of dividing track A into constant intervals of one second in the same manner as in the above described application of a CD-ROM. That is to say, the length of the segments may vary according to the edit policy at the time of recording. Accordingly, a medium that can be randomly accessed, such as a CD-ROM, may have sections created by separating the medium into constant time intervals.

The present invention is not limited to the above described embodiments but rather a variety of modifications are possible.

The present invention can be applied to a system, for example, that connects to a video disk or a video tape. That is to say, a soundtrack of a movie is placed on track A spoken at regular speed (speed of a native speaker); speech sound spoken slowly

is placed on track B; and Japanese explanation is placed on track C. Thus, the reproduction of track A can be stopped at a place that cannot be understood and slightly moved back (at this time, the image can be held still and at the same time the system can be practically used) so that track B can be heard and if it cannot still be understood track C can be used. Here, the speech sound may be synchronized with the images only at the time track A is reproduced.

In addition, the present invention can be used on a personal computer or the like. That is to say, a high level application becomes possible by combining with a system connecting a CAI (Computer Aided Instruction) or a programmable high level system such as a CDI (Compact Disk Interactive). In addition, it is also possible to use the memory device of a personal computer in place of the above described CD or DAT in order to implement the present invention.

Furthermore, the present invention can be applied to the study of recitation of Chinese poems, the study of law, in addition to the study of English conversation. Moreover, data sequences are not limited to three types or four types, but rather may be of two types or no less than five types.

[Effects of the Invention]

As described above in detail, a data recording medium

according to the present invention can record speech sound data, and the like, having appropriate contents desired by experts and novices and, therefore, a data recording medium that can satisfy both experts and novices can be gained.

In addition, according to a data reproduction system in accordance with the present invention, speech sound data, and the like, having appropriate contents for experts and novices can be reproduced by using a recording medium according to the present invention and, therefore, a data reproduction system that can satisfy both experts and the novices can be gained.

4. BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is an explanatory view showing data sequences A, B, and C and the recorded contents thereof when a recording medium of the present invention is applied for the self-study of English conversation; Fig. 2 is an explanatory view showing the relationship between a period of time and a capacity of each segment in the example shown in Fig. 1; Fig. 3 is an explanatory view showing one example of a directory of the example shown in Fig. 1 and Fig. 2; Fig. 4 is a block diagram showing the configuration of one example of a reproduction unit; Fig. 5(a) and Fig. 5(b) are diagrams showing memory allocations in ROM 6 and in RAM 7; Fig. 6 is an explanatory

view showing handset 8 and earphone 13 showing Fig. 4; Fig. 7 is a perspective view of the reproduction unit shown in Fig. 4; Fig. 8 is a view showing the condition where the handset is hand held; Fig. 9 is a sequence diagram showing the effects; Fig. 10 is an explanatory view showing a directory according to the second example of the present invention; Fig. 11 is a view showing divided segments of data sequences A, B and C according to the directory of Fig. 10; Fig. 12 is a view showing the relationship between a value of the second and third bits and a segment of attribute C of 8 bits; Fig. 13 is an explanatory view showing a directory according to the third example of the present invention; and Fig. 14 is a frontal view of a handset used in the reproduction unit according to this example.

1 ... medium, 2 ... reproduction mechanism, 8 ... handset, 13 ... earphone, 21 ... segment display part, 22 ... FF switch, 23A ... A instruction switch, 23B ... B instruction switch, 23C ... C instruction switch, 24 ... STOP button, 25 ... REV button, 41 ... text display part, and 42 ... text display button

Fig.1

Recorded Contents	
A	<p>"It's not much of a problem. I'd second that."</p> <p>621 622 623 624 625 626 627</p>
B	<p>"It is not much of a problem. I would second that."</p> <p>621 622 623 624 625 626 627</p>
C	(621~624) "'It' indicates the previous 'something' and 'not much of a problem' is an idiom meaning 'no problems' or 'no worries'."
	(625~627) Next, 'I'd' is an abbreviation of 'I would' and 'I could' can also be abbreviated to 'I'd' in the same manner. 'Second' means to 'support' or to 'agree with' and 'I'd second that' means 'I agree with that'."

Fig.11

Data Sequence	
Divided Segments	
A	<p>It's not much of a problem. I'd second that.</p> <p>621 622 623 624 625 626 627</p>
B	<p>It is not much of a problem. I would second that.</p> <p>621-1 621-2 622 623-1 623-2 623-3 624 625-1 625-2 626 627</p>
C	<p>621-1 621-2 622 623-1 623-2 623-3 624 625-1 625-2 626 627</p> <p>□ □ □ □ □ □ □ □ □ □ □</p>

Fig.12

Second or Third Bit	Segment
0	621
1	621-1
1	621-2
2	621-1
3	621-2
0	622
1	622
2	622
0	623
1	623-1
1	623-2
1	623-3
2	623-1
2	623-2

Fig.2

		6 2 1	6 2 2	6 2 3	6 2 4	6 2 5	6 2 6	6 2 7	
Time (second)	A	0. 2	0. 1	0. 4	0. 3	0. 1	0. 3	0. 4	1. 8
	B	0. 9	0. 2	1. 2	0. 6	0. 8	0. 6	0. 7	5. 0
	C	1 2				2 5			3 7
Capacity (KB)	A	1. 2	0. 6	2. 4	1. 8	0. 6	1. 8	2. 4	10. 8
	B	5. 4	1. 2	7. 2	3. 6	4. 8	3. 6	4. 2	30. 0
	C	7 2				1 5 0			2 2 2

Fig.3

Data Sequence	Segment	Category	Position Data			Start Byte	Length (B Y T E)
	S S	C	M	S	B	S B	L L L
A	6 2 1	0	0	1 1	3	8 2 6	1, 2 0 0
B		6 4	0	1 1	3	2, 0 2 6	5, 4 0 0
C		1 2 8	0	1 1	6	1, 2 8 2	7 2, 0 0 0
A	6 2 2	0	0	1 1	4 1	1, 6 0 2	6 0 0
B		6 4	0	1 1	4 2	1 5 4	1, 2 0 0
C		1 2 9	0	1 1	6	1, 2 8 2	7 2, 0 0 0
A	6 2 3	0	0	1 1	4 2	1, 3 5 4	2, 4 0 0
B		6 4	0	1 1	4 3	1, 7 0 6	7, 2 0 0
C		1 2 9	0	1 1	6	1, 2 8 2	7 2, 0 0 0
A	6 2 4	0	0	1 1	4 7	7 1 4	1, 8 0 0
B		6 4	0	1 1	4 8	4 6 6	3, 6 0 0
C		1 2 9	0	1 1	6	1, 2 8 2	7 2, 0 0 0
A	6 2 5	0	0	1 1	4 9	2, 0 1 8	6 0 0
B		6 4	0	1 1	5 0	5 7 0	4, 8 0 0
C		1 2 8	0	1 1	5 2	1, 2 7 4	1 5 0, 0 0 0
A	6 2 6	0	0	1 2	5 0	1, 7 7 0	1, 8 0 0
B		6 4	0	1 2	5 1	1, 5 2 2	3, 6 0 0
C		1 2 9	0	1 1	5 2	1, 2 7 4	1 5 0, 0 0 0
A	6 2 7	0	0	1 2	5 3	1, 0 2 6	2, 4 0 0
B		6 4	0	1 2	5 4	1, 3 7 8	4, 2 0 0
C		1 2 9	0	1 1	5 2	1, 2 7 4	1 5 0, 0 0 0

Fig.4

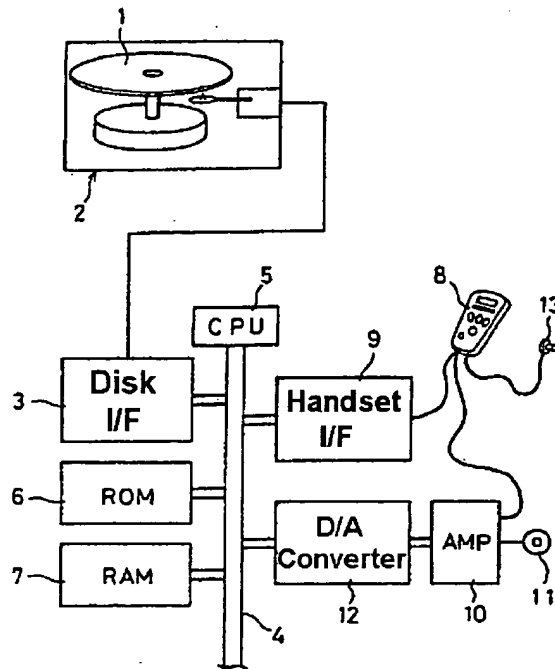


Fig.5

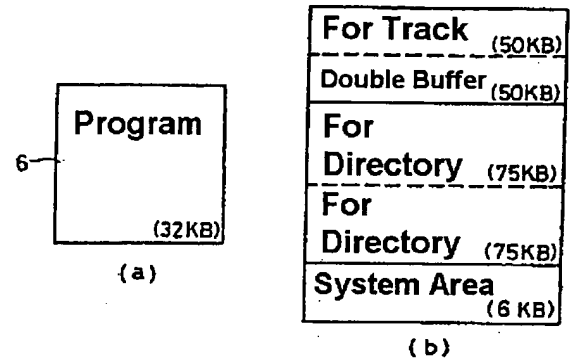


Fig.6

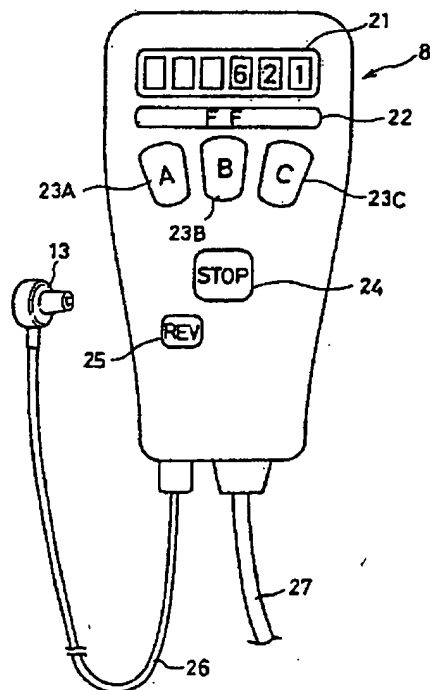


Fig.7

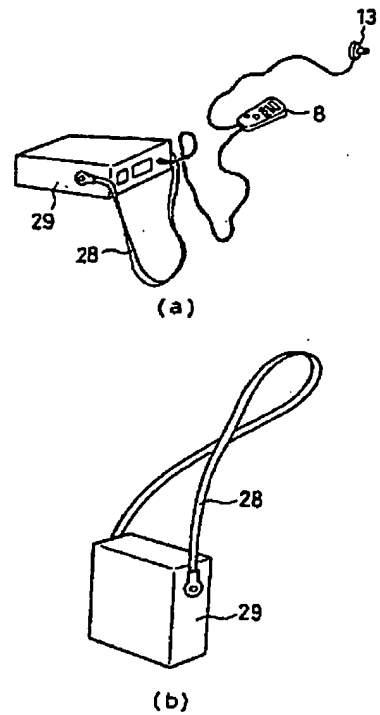


Fig.8

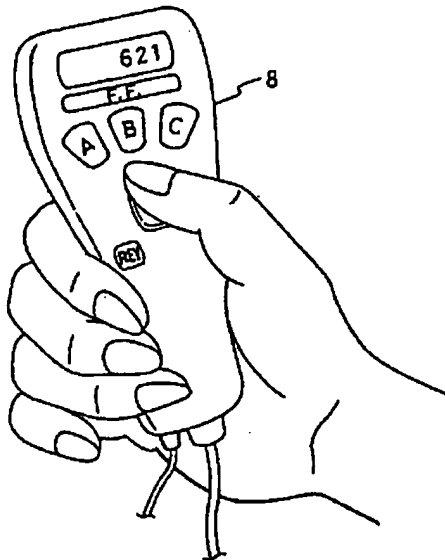


Fig.10

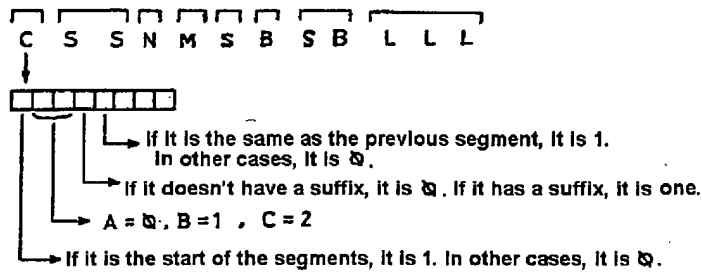


Fig.13

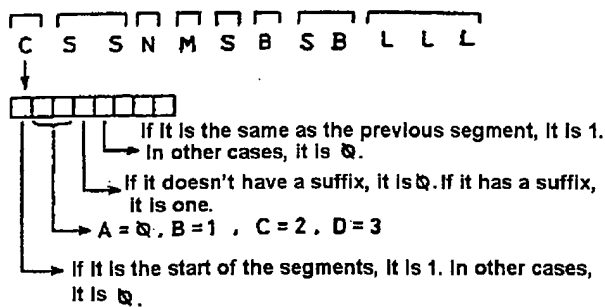


Fig.9(a)

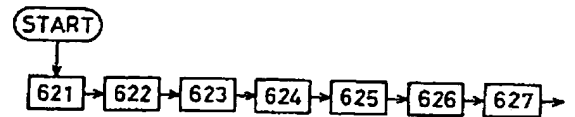


Fig.9(b)

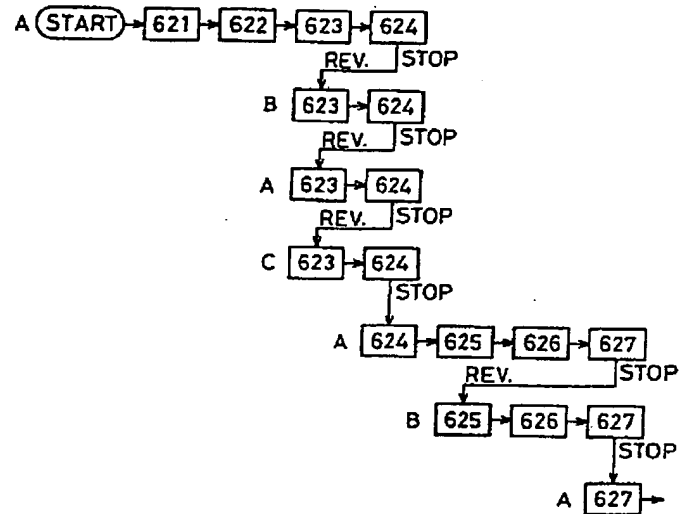


Fig.14

